Delayed, Not Deterred? Tuition Fees and University Enrolment

Luke Walsh*
Department of Economics
University of Liverpool

Abstract

Reducing subsidies to higher education raises concerns about people's willingness to invest in human capital, yet little is known about how students respond to price in settings where credit constraints are absent. I exploit the UK's 2012 tuition fee reform, which nearly tripled the price cap on university tuition while retaining universal income-contingent loans, using a difference-in-discontinuities design built on England's September school-entry cutoff. Entry at age 18 fell by around four percentage points, but cumulative entry by age 23 shows a much more modest 0.6-7pp decline: the enrolment response was mostly a change in timing, not a permanent withdrawal from higher education. The reform durably changed where students enrolled. Entry into more selective institutions rose across the selectivity distribution, implying a permanent reallocation away from less selective institutions within a cohort of unchanged size, and this shift is present among students unaffected by concurrent supply-side policy. I find no effects on degree attainment or subject choice, no heterogeneity by socioeconomic background, gender, or prior attainment, and an employment-age profile that independently confirms the retiming dynamics. These findings show that a large reduction in state subsidies, when paired with a progressive loan system, reshaped the timing and destination of human capital investment with only modest reductions in levels and without widening inequality.

IMPORTANT NOTE ON PRELIMINARY RESULTS

The results presented in the tables and figures in this draft are affected by a known data error involving a missing year of observations. The underlying data have now been corrected, but revised outputs are awaiting release from the Secure Research Service (SRS). **Where the tables or figures conflict with the discussion in the text, readers should defer to the text.** Numerical estimates reported in the text have been suppressed where necessary in accordance with SRS statistical disclosure control requirements.
The tables and figures should therefore be treated as provisional and will be replaced once the corrected outputs have been released.

*Corresponding Author: l.walsh7@liverpool.ac.uk.

1 Introduction

The cost of higher education plays a central role in human capital investment, shaping both individual decisions to pursue further study and governments' willingness to subsidise that investment. Classic human capital models predict that individuals invest in higher education when the present value of future returns, net of costs, exceeds the value of entering the labour market directly (Becker 1964). An exogenous increase in the cost of university should therefore reduce demand at the margin. Yet price changes can operate on margins that a simple participation model obscures: students may respond by changing *when* they enrol, or *where*, rather than *whether*. Distinguishing these margins matters for policy, because a fee increase that permanently prices students out of higher education has very different welfare implications from one that reshuffles the timing and destination of an unchanged level of investment.

A widely cited rationale for subsidising higher education is the presence of credit constraints: individuals with strong academic potential but no collateral may be unable to finance tuition even when expected returns are high (Black et al. 2023; Lochner and Monge-Naranjo 2011). The UK's 2012 tuition fee reform provides a setting in which this friction can be largely set aside. The reform nearly tripled the cap on annual tuition fees, from £3,375 to £9,000, but because all students retained access to government-backed income-contingent loans, no student faced an upfront payment and repayment remained contingent on future earnings. Fee variation in this system operates through expected lifetime returns rather than liquidity, providing a clean test of how students respond to the price of higher education when the under-provision of capital is not a factor.

I estimate the causal effect of the reform by exploiting its alignment with England's school-entry cutoff. Children born on or after September 1st enter school one academic year later than those born just before, generating a well-documented discontinuity in educational outcomes at the cutoff. Because students enter university in the academic year following their 18th birthday, those

born in September 1993 were the first cohort fully exposed to the higher fees, while those born in August 1993 were not. My difference-in-discontinuities design tests whether the stable birthday premium in enrolment changed for the treated cohort, and I conduct inference by comparing the treated-year estimate against the empirical distribution of placebo discontinuities in unaffected cohorts (Conley and Taber 2011), an approach suited to a design with a single treated cohort, and one whose validity is transparent: the placebo estimates are tight around zero, while the treated-year estimates lie far outside their distribution.

The central result is a decomposition of the enrolment response into two distinct estimands. The *total* effect of the reform and its announcement on entry at age 18 is a decline of around four percentage points. But this measure conflates permanent deterrence with retiming: the reform was announced two years in advance, giving the final pre-reform cohort a strong incentive to forgo deferral and enter in 2011, and giving treated students facing the new regime weaker incentives to enter on time. I document this anticipation by showing that deferred entry declines in 2011 and address it with cumulative outcomes that are immune to retiming by construction: a student who enters at 19 rather than 18 counts identically in entry by age 21. The *permanent* effect shrinks to about 0.6-0.7pp, which should be viewed cautiously as a floor due to control contamination mechanics, but appears stable nonetheless.

The reform did, however, permanently change where students enrolled. Using ever-entered measures that are likewise robust to retiming, I find that entry into more selective institutions rose after the reform: entry into institutions above the median, in the top quartile, top quintile, and top decile of selectivity rose by [X], [X], [X], and [X] percentage points respectively. Combined with the small permanent effect on total entry, this implies that a cohort only marginally smaller in the end attended a markedly different set of institutions, with the rise in selective entry mirrored by a larger fall in entry to less selective institutions. The fee increase operated principally by reallocating students across the quality distribution of higher education rather than by pricing them out of it.

The remaining results support this interpretation and speak to the reform's costs. I find no effects on degree failure or non-completion; the students who delayed or traded up did not do observably worse. I find no heterogeneity by free school meal eligibility, gender, neighbourhood university participation, or prior attainment, consistent with the progressive design of the income-contingent loan system insulating disadvantaged students from the fee increase. And the labour market provides independent confirmation of the enrolment dynamics: the treated cohort exhibits higher employment at ages 19–21, when delayed entrants are working before enrolment, and lower employment at ages 23–24, when they remain in study while their counterfactual on-time peers have graduated. By age 25, I find no evidence that the reform damaged earnings or employment, alongside suggestive evidence of movement into higher-paying fields within the treated cohort.

These findings reconcile an apparent tension in the existing evidence. Studies of the 2012 reform report small or negligible effects on overall participation (Azmat and Simion 2021), while the raw entry data show a sharp dip in 2012. Both observations are correct: the dip was real but temporary, an artefact of announcement-induced retiming around the reform, while the permanent participation effect was close to zero. My contribution is the decomposition itself, separating the total effect of a pre-announced price change from its permanent component, together with the destination margin, which snapshot designs centred on a single entry year cannot observe. Prior UK studies, constrained by application-year or survey data, could measure neither the timing dynamics nor where the cohort ultimately enrolled (Dearden et al. 2011; Sá 2014; Azmat and Simion 2021).

Methodologically, the paper leverages population-wide administrative data linking the National Pupil Database (NPD), Higher Education Statistics Agency (HESA) records, and HMRC tax data. This Longitudinal Education Outcomes (LEO) dataset allows me to track entire cohorts from secondary school through university entry at any age and into the labour market, which is precisely what the retiming decomposition requires. The design also contributes a template for

inference in difference-in-discontinuities settings with a single treated cohort, where conventional clustered standard errors are unreliable and the placebo distribution of pre-reform discontinuities provides the honest benchmark.

The results carry a clear policy message. In a system with universal income-contingent loans, a large increase in the private cost of university produced only a modest permanent reduction in participation and no widening of participation gaps; its dominant effects were on the timing of entry, temporarily, and the destination of enrolment, permanently. The margin on which students principally respond to price in this setting is quality rather than quantity. This reframes the debate over fee levels from one about access towards one about the allocation of students across institutions.

The remainder of the paper is structured as follows. Section 2 reviews the related literature. Section 3 provides institutional background on the UK higher education system and the 2012 reform. Section 4 describes the data sources and sample construction. Section 5 outlines the empirical strategy. Section 6 presents the main findings, and Section 7 reports robustness checks and additional analyses. Section 8 concludes.

2 Related Literature

Human capital theory predicts that students weigh the costs of university against its expected long-term returns, and that an increase in price should reduce investment at the margin (Becker 1964; Mincer 1974). But this prediction concerns the *level* of investment, and empirical work on tuition prices has largely inherited that focus, asking whether fees reduce participation. A price change can instead operate on margins that a participation regression cannot see: the timing of entry, particularly when the change is announced in advance, and the choice of institution conditional on entering. The contribution of this paper is to measure all three margins within a single design, and its central finding is that the response to the 2012 reform was concentrated

almost entirely on the timing and destination margins rather than the participation margin.

A large literature examines the participation margin directly. In North America, where tuition varies across states, provinces, and institutions, enrolment falls modestly when sticker prices rise (Kane 1995; Neill 2009; Hemelt and Marcotte 2011); see Deming and Dynarski (2009) for a review. The closest international analogue to the English experience is Germany, where the staggered introduction of tuition fees across states reduced enrolment intentions and transitions into university by small amounts (Hübner 2012; Bruckmeier and Wigger 2014), and where the subsequent abolition of those fees provides symmetric evidence of modest price sensitivity. These studies consistently find small participation elasticities, but their designs measure enrolment in a fixed window around the price change and therefore cannot distinguish permanent deterrence from delay which is precisely the distinction this paper is built to make.

A related American literature studies financial aid rather than posted prices, finding that need-based grants raise enrolment and persistence among low-income students (Dynarski 2003; Bettinger 2004; Castleman and Long 2016), with downstream benefits for completion and earnings that can exceed their fiscal cost (Denning et al. 2019; Black et al. 2023). This work is conceptually distinct from mine: aid operates by relaxing liquidity constraints for students who cannot borrow against future earnings, whereas a fee change under a universal income-contingent loan system alters the expected lifetime return to university without any liquidity effect. The English setting is therefore unusually clean for isolating the price mechanism: because no student faces an upfront payment and repayment is contingent on earnings, the fee increase operates through expected returns alone, and the debt-aversion channel documented elsewhere (Callender and Jackson 2005; Field 2009) is the residual behavioural concern rather than a binding financial one.

The UK literature on the 2012 reform itself has produced an apparent puzzle. Earlier fee increases reduced participation modestly (Dearden et al. 2011; Sá 2014), and evaluations of the 2012 reform conclude that it left overall enrolment and the socioeconomic composition of entrants essentially unchanged (Murphy et al. 2019; Azmat and Simion 2021), with Azmat and

Simion (2021) additionally reporting minimal effects on early labour market outcomes. Yet the raw entry data show a sharp, visible drop in 2012. My results reconcile these observations: the 2012 dip was real but temporary, generated by announcement-induced retiming. The final pre-reform cohort abandoned deferral to enter in 2011, and treated students entered late. While the permanent participation effect was small, exactly as the consensus holds. Where I depart from the existing UK evidence is on the destination margin. Because prior studies rely on application-year snapshots or the DLHE survey, they observe neither entry at later ages nor where a cohort ultimately enrolled; using linked administrative data that track entry at any age, I show the reform permanently reallocated students toward more selective institutions.

The retiming result connects to a broader methodological point about pre-announced policies. When agents can shift the timing of a decision, comparisons of outcomes immediately before and after implementation confound the policy's permanent effect with intertemporal substitution around it, and pre-period observations are contaminated as controls (Malani and Reif 2015). The 2012 reform, announced two years before implementation with a fee schedule tied to entry date, created textbook conditions for such substitution. I treat this not as a nuisance but as an object of study: I document the anticipation response directly, define separate estimands for the total and permanent effects, and use cumulative entry outcomes that are immune to retiming by construction.

The identifying variation rests on the relative age effect which is the well-documented advantage of autumn-born children within academic cohorts (Russell and Startup 1986; McPhillips and Jordan-Black 2009; Crawford et al. 2013, 2014; Black et al. 2011), and on the stability of that premium across adjacent cohorts, which I verify directly. Season-of-birth selection, the classic threat to birthdate designs (Buckles and Hungerman 2013), is differenced out to the extent it is stable across cohorts, and I show balance on predetermined characteristics at the cutoff for every cohort in the sample.

3 Institutional Background

In the decades following WWII, the funding of students into higher education was largely administered by local authorities, an approach formalised by the Education Act (1962). The Act mandated that Local Education Authorities pay tuition fees and provide means-tested maintenance grants for eligible full-time university students (UK Parliament 1962), making undergraduate education effectively free at the point of use across the UK. The economic rationale was summarised in the Robbins Report of 1963, which recommended a substantial expansion of the university system on the grounds that higher education generated considerable positive social and financial externalities, such that growth in the economy and in tax receipts would deliver a positive return on public investment (Committee on Higher Education 1963).

This assumption came under strain in the 1990s. Student numbers had grown rapidly while funding per pupil declined by 40% between the early 1980s and 1997, putting considerable pressure on university budgets (Dearden et al. 2011). At the same time, the case for higher education paying for itself at the exchequer level had begun to wane: the Dearing Report (1997) argued that a cost-sharing approach would be equitable since graduates were the primary beneficiaries of university education (National Committee of Inquiry into Higher Education 1997). Subsequent evidence supports this reassessment: Britton et al. (2020) estimate that while average lifetime returns to the exchequer are positive net of subsidies, these returns are driven predominantly by the highest earners, and total returns are negative for up to 30% of graduates.

3.1 Post-1998 Reforms

Following the Dearing Report, the Teaching and Higher Education Act (1998) introduced annual tuition fees of £1,000, paid upfront by students, while maintenance grants were replaced with income-contingent maintenance loans (UK-Government 1998; Hillman 2013). Exemptions applied to students from low-income families, and following devolution in 1999 the UK nations

began to diverge: Scotland introduced a graduate endowment in lieu of upfront fees and by 2007 had abolished tuition fees entirely for Scottish-domiciled undergraduates at Scottish institutions (Sá 2014).

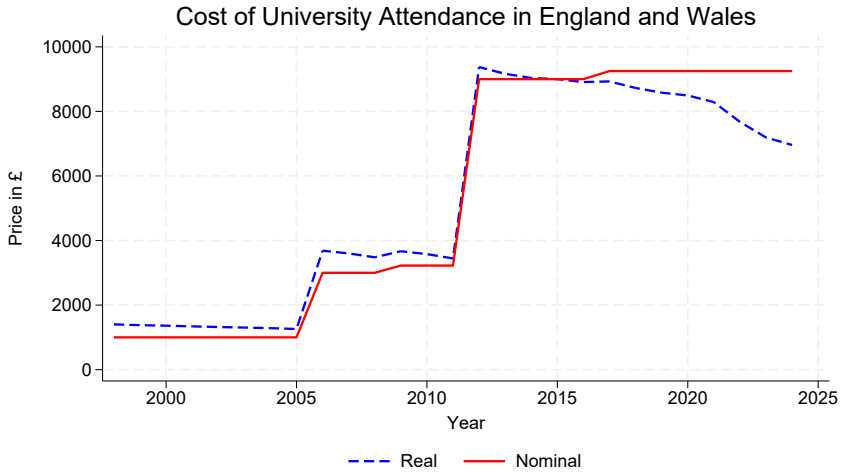
By the early 2000s the £1,000 cap was yielding limited revenue, and the requirement that fees be paid upfront was widely seen as a barrier to access. The Higher Education Act 2004 raised the cap to £3,000 per year from the 2006–07 academic year and, crucially, ended upfront payment altogether (UK Parliament 2004). Fees would instead be covered by government-backed loans, repaid after graduation at 9% of earnings above a £15,000 threshold, with any outstanding balance written off after 25 years. University thereby became free at the point of use once more, with higher fees operating as a reduction in expected lifetime income rather than a liquidity burden. This institutional detail distinguishes the UK setting from the US, where the policy debate centres on the under-allocation of capital to students who cannot collateralise their future earnings, and it is what allows this paper to interpret responses to the 2012 reform as responses to price rather than to credit constraints.

3.2 The 2012 Reform and Its Announcement

By 2010, recession and the declining real value of the capped fee had renewed the pressure on university finances. The independent Browne Review, published in October 2010 and accepted by the government in November of the same year, recommended a market-driven increase in the fee cap to £9,000 per year, effective for students entering university from the 2012–13 academic year (Independent Review of Higher Education Funding and Student Finance 2010). In theory variable fees were intended to generate price competition; in practice almost all universities charged the maximum immediately upon implementation (House of Commons Library 2018). The reform also shifted the repayment threshold from just below £16,000 to £21,000, extended the repayment period from 25 to 30 years, and introduced a higher, income-contingent interest rate (House of Commons Library 2018). The net effect was a more progressive schedule: the lowest-earning

graduates could expect to pay less per month than under the previous system (Britton et al. 2017). Figure 1 summarises the evolution of the fee cap, with the 2012 increase standing as a clear outlier against the earlier reforms.

Figure 1



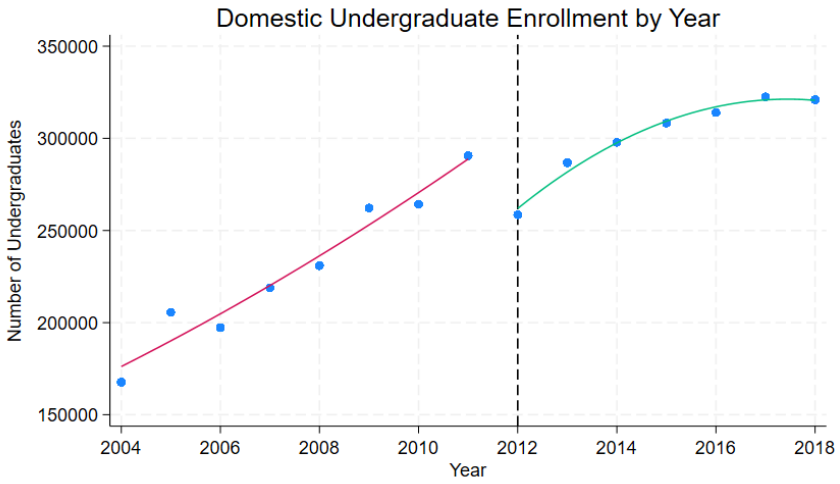
Note: Figure shows the maximum annual tuition fee chargeable to home undergraduate students in England, 1998–2024, in nominal and real (2024) prices. The 1998, 2006, and 2012 discontinuities correspond to the Teaching and Higher Education Act (1998), the Higher Education Act (2004), and the post-Browne reform respectively.

Two features of the reform's implementation are central to the empirical design. First, the fee regime a student faced was determined by their year of entry, not their year of birth: anyone commencing study before September 2012 paid the old fees for the duration of their course, while anyone commencing from September 2012 onwards paid the new fees. Second, the reform was announced almost two years before it took effect. Together these created sharp incentives to retime entry around the reform: students in the final pre-reform cohort could lock in the lower fees by entering in 2011 rather than deferring, while students in the treated cohort, already facing the higher fees whenever they entered, lost nothing by delaying. I document in Section 6 that deferred entry collapsed in 2011, precisely as these incentives predict. Rather than treating anticipation as a nuisance, the analysis defines estimands that separate this retiming response from the reform's permanent effect on participation.

Figure 2 displays the raw trend in domestic undergraduate enrolment over the sample period,

and exhibits exactly the pattern the incentives above would generate: steady year-on-year growth followed by a sharp contraction in 2012, and a recovery in subsequent years. Taken at face value, the 2012 dip suggests a substantial deterrent effect of the fee increase; the central task of this paper is to decompose how much of it reflects students permanently priced out of higher education and how much reflects entry shifted across years.

Figure 2



Note: Figure shows the number of domestic first-year undergraduate entrants at English institutions by academic year of entry, 2004–2018. The dashed line marks the 2012–13 academic year, the first cohort subject to the £9,000 fee cap. Fitted trends are shown separately for the pre- and post-reform periods.

3.3 Concurrent Policies

Three contemporaneous policy changes could in principle confound the analysis, and each is addressed directly. First, and most importantly, the government exempted students achieving grades of AAB or higher at A-level from student number controls in the 2012–13 academic year, extended the exemption to ABB+ in 2013–14, and abolished number controls entirely from 2015–16. This gave selective institutions the freedom to expand recruitment of high-attaining students at precisely the moment the fee reform took effect, and is therefore a candidate explanation for any shift in enrolment toward selective institutions. Section 6 shows that the reallocation I document is present among students with attainment *below* the AAB threshold, who were unaffected

by the exemption. Second, the participation age was raised, requiring young people in England to remain in education or training until age 17 from 2013 and 18 from 2015. This affects cohorts entering university from 2015 onwards and is handled by the choice of estimation window. Third, maintenance support evolved over the period, but these changes were incremental and applied smoothly across cohorts; they are differenced out by the design, which identifies effects only from discontinuous changes specific to the treated cohort (House of Commons Library 2018).

4 Data and Descriptive Statistics

This paper uses data from the Longitudinal Education Outcomes (LEO) dataset. This administrative panel links school census records from the National Pupil Database (NPD) on student characteristics to Higher Education Statistics Agency (HESA) records of university enrolment and outcomes, matched further to HMRC administrative data on earnings, employment, and benefits. The result is a longitudinal panel covering over 30 million people across a 22-year period (1998–2020), from which I construct cohorts of English-domiciled, state-educated pupils defined by academic year of birth. Because the panel follows individuals rather than entry classes, I observe university entry at any age, not merely entry in the year following a student's 18th birthday, a feature that permits the decomposition of the enrolment response into its retiming and permanent components.

For each individual I construct a set of entry indicators: entry in the academic year following their 18th birthday, cumulative entry within two through five years of that date, and an ever-entered indicator over the observed horizon. Cumulative entry by age 23 for the treated cohort requires HESA coverage through the 2017–18 academic year, which the panel provides. Institution selectivity is measured by ranking institutions according to the (mean) average A-level points of their entering students for the 2011 cohort, from which I define indicators for entry into institutions above the mean and in the top quartile, quintile, and decile of the distribution;

each is constructed both conditional on entry and unconditionally. Labour market outcomes are drawn from the HMRC paye as you earn (PAYE) data, and comprise employment and pay observed at each age, as well as the industry of employment, which I rank by median pay using a cross-section of the 2015 tax year.[1]

Personal characteristics contained within the NPD are used for heterogeneity analysis and balance testing. Family income is not recorded, so eligibility for free school meals (FSM) in year 11, a means-tested programme for children from low-income families, is used as a proxy for childhood poverty. Neighbourhood disadvantage is measured by residence in a local area with below-mean university participation (LOPAR). Prior attainment is measured by standardised total KS5 point scores, from which I also construct an indicator for achieving grades of AAB or higher at A-level which is the threshold at which students were exempted from student number controls in 2012–13 (Section 3.3).[2]

Table 1 presents summary statistics for first-year university entrants immediately before and after the reform. The composition of entrants remained broadly stable between the 2011 and 2012 cohorts: the share of female students was unchanged, and the shares eligible for FSM and from low-participation areas rose marginally. Two margins moved. Entry in the year following students' 18th birthdays fell by 2.3 percentage points, and the share of entrants attending an above-average institution rose from 45 to 48 percent. These raw movements preview the paper's two central findings: a fall in on-time entry alongside a shift toward more selective institutions. Though neither difference can be interpreted causally without the design set out in Section 5.

[1] LEO does not contain occupation codes, so industry selection may not map one-for-one into occupational upgrading. The ranking of industries is highly robust to the choice of reference year.

[2] In accordance with the disclosure rules governing LEO, all reported counts are rounded to the nearest five.

Table 1: University Entrant Characteristics Before and After the 2012 Fee Reform

	2011	**2012**
Entry into university at 18 (%)	24.8	22.5
Female (%)	55.7	55.7
Eligible for FSM (%)	9.0	9.6
From low university-share area (%)	56.0	56.8
AAB or higher at A-level (%)	[X]	[X]
Attending above-average institution (%)	44.8	47.9

Notes: This table reports the share of first-year university entrants by key demographic and background characteristics in 2011 (pre-reform) and 2012 (post-reform). FSM = Free School Meals. "Low university-share area" indicates local areas with below-average university participation in 2011. "Attending above-average institution" refers to institutions ranked above the median in terms of average A-level points of entrants. Counts rounded to the nearest five per LEO disclosure rules.

5 Methodology

This paper asks whether a sharp increase in the cost of university affected students' decisions about whether, when, and where to attend, and how they subsequently fared in the labour market. Estimating these effects is difficult: tuition fees typically evolve alongside other policy changes, broader macroeconomic shocks may confound simple comparisons, and students facing a weak labour market may enrol in university as a form of strategic delay. A simple comparison of adjacent cohorts cannot isolate the effect of the fee change from these influences.

To overcome this, I exploit a natural experiment created by the structure of England's school-entry system and the timing of the 2012 reform. In England, children are placed into academic cohorts based on their birthdate, with a strict cutoff on 1 September. Those born just before the cutoff (e.g. in August) enter school a full year earlier than those born just after (e.g. in September), even though they are only days apart in age. This generates sharp and well-documented differences in long-term outcomes, including university attendance and earnings, between children born on either side of the cutoff, commonly referred to as the relative age effect (Russell and Startup 1986; McPhillips and Jordan-Black 2009; Campbell 2013; Crawford et al. 2013, 2014). I show evidence of this premium in Figure 3: September-born individuals are significantly more

likely to enrol in university, with a steady decline in each subsequent birth month. Across my sample, students born in August are more than 3 percentage points less likely to enrol than those born the following September.

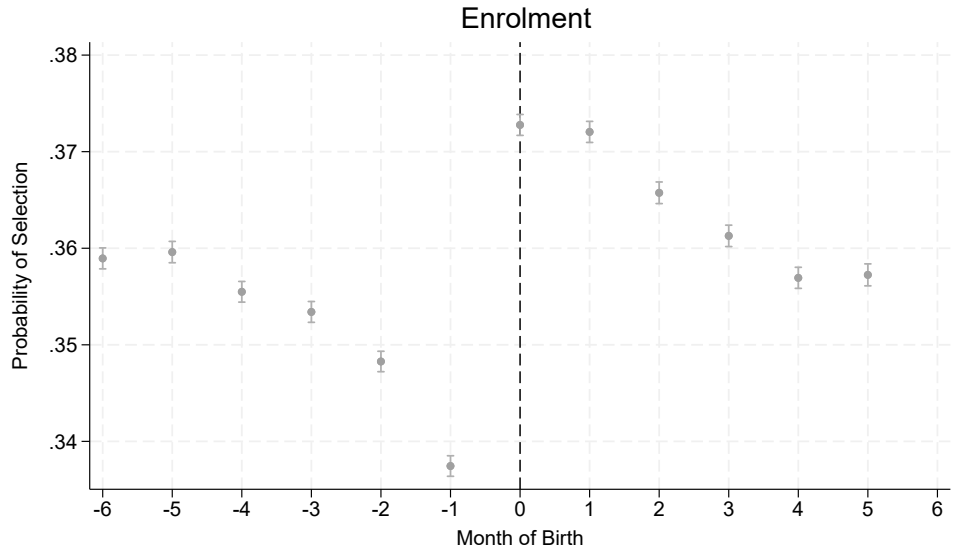


Figure 3: Probability of university enrolment by month of birth, centred at 0 for September-born students, illustrating the discontinuity around the academic year cutoff.

The 2012 reform raised the maximum fee from £3,375 to £9,000 for students entering higher education from September 2012 onwards. Because students typically enter university in the academic year following their 18th birthday, those born in September 1993 were the first cohort whose on-time entry fell under the higher fees, while those born in August 1993 could enter on time under the old regime. The policy's incidence therefore aligns with the existing school-entry discontinuity: within a single pair of adjacent birth months, students days apart in age faced different fee regimes for on-time entry.

My identification strategy exploits this structure by examining whether the relative age effect deviates from its stable pattern in the treated cohort. If the fee increase had no effect, the size of the September–August discontinuity should be the same in the treated cohort as in every other cohort. A change in the discontinuity that occurs only in the cohort whose on-time entry was treated can therefore be interpreted as the causal effect of the policy. This difference-in-discontinuities

design leverages two sources of plausibly exogenous variation: the sharp birthdate cutoff, and the reform that changed fees for a specific cohort. I estimate:

$$Y_{i,m,t} = \alpha + \tau\,(D_i \cdot \text{Fee}_i) + \gamma D_i + \delta\,\text{Fee}_i + \beta_1 X_i + \lambda_t + \epsilon_{i,m,t}$$

where $Y_{i,m,t}$ is the outcome of individual i with month of birth m and academic year of birth t: an entry indicator, an institution-tier indicator, or a labour market outcome. D_i is the discontinuity indicator, equal to 1 if individual i was born in September (the first month of the academic year) and 0 if born in August. Fee_i is the treatment-cohort indicator, equal to 1 for individuals whose on-time entry year falls under the higher fee regime and 0 otherwise. X_i is the month-of-birth running variable, λ_t are academic-year-of-birth fixed effects, and $\epsilon_{i,m,t}$ is the error term. The coefficient of interest, τ, is the difference-in-discontinuities estimator: the change in the September–August gap that occurs only in the treated cohort. Conceptually, I measure the gap between autumn-born and summer-born individuals in every cohort, and test whether that gap changes significantly in the cohort exposed to higher fees. Standard errors are clustered at the cohort level, though as set out in Section 5.2, inference does not rest on them.

Estimates of the discontinuity change are computed for every pre-reform cohort in the sample as well as the treated cohort, and every results figure displays them together. The pre-reform estimates serve as placebo tests: cohorts facing no fee change should exhibit no change in the relative age effect. These placebo estimates are computed excluding the treated cohort, since its inclusion would bias the estimated birthday premium and thereby inflate the placebos. Because the treated cohort is omitted from those estimations, the final pre-reform cohort cannot be separately identified due to collinearity with the year-of-birth fixed effects, and does not appear as a placebo in the central specification.

5.1 Estimands: Total, Retiming, and Permanent Effects

The reform's two-year announcement window, combined with the rule tying fees to entry year rather than birth year (Section 3), means that a single enrolment outcome cannot summarise the policy's effect. Entry in the year following a student's 18th birthday responds both to permanent decisions not to attend and to decisions to shift entry across years: control-cohort students had a strong incentive to forgo deferral and enter in 2011, and treated students, facing the higher fees whenever they entered, lost nothing by delaying. I therefore define two estimands. The *total* effect, τ^{18}, is the effect on entry at age 18, and captures the full behavioural response to the reform and its announcement, retiming included. The *permanent* effect, τ^{a}, is the effect on cumulative entry within a years, for horizons of two through five years. Cumulative outcomes are immune to retiming by construction: a student who enters at 19 rather than 18 counts identically in entry by 21, on either side of the cutoff. The difference between the total and long-horizon estimates measures the retiming component.

One complication deserves direct treatment. Because fees followed entry year, a control-cohort student who delayed entry beyond 2011 paid the higher fees: the control group's late-entry margin is itself treated. This contamination cannot generate a spurious permanent effect; it attenuates the estimate toward zero, because deterred late entry in the control cohort absorbs part of the treated cohort's deterred entry. The long-horizon estimates are therefore contaminated; this can be seen directly in the slope of the fourth panel of Figure 4. Estimates of ever enrolment should therefore be interpreted as a lower bound, though it's worth noting that the extent of contamination is very small within three years of treatment.

5.2 Inference

Inference in this design requires care, and conventional tools are poorly suited to it. With a single treated cohort and a small number of control cohorts, cluster-robust standard errors are

known to be unreliable (Cameron et al. 2008). Nor do the standard remedies for few clusters apply: wild cluster bootstrap procedures require variation within multiple treated clusters and are uninformative when treatment is confined to one (MacKinnon and Webb 2018), and an exact permutation test across the seven placebo cohorts in the sample cannot deliver a p-value below one eighth, too coarse to certify significance at conventional thresholds however extreme the treated estimate.

The primary basis for inference throughout the paper is therefore the placebo distribution itself, in the spirit of Conley and Taber (2011); Hagemann (2019); MacKinnon and Webb (2020). Every results figure reports the estimated discontinuity change for each untreated cohort alongside the treated cohort, and the evidential weight of a result derives from two features that the reader can verify directly: the placebo estimates are tight around zero, demonstrating that the relative age effect is stable in the absence of a fee change, and the treated estimate lies far outside the range the placebos span. This comparison uses no estimated standard errors. Its logic is that of randomisation inference: the placebo estimates trace out the distribution of discontinuity changes that arise from noise and minor cohort shocks alone, and a treated estimate several times the largest placebo cannot plausibly be a draw from that distribution, even though the small number of placebos precludes attaching a precise p-value to the statement. Where conventional standard errors appear in tables, I report them for transparency and adopt significance at the 1 percent level as the reporting threshold, a deliberately conservative convention; no claim in the paper rests on a conventional standard error alone, and no result is asserted whose treated estimate sits inside the span of its placebos.

5.3 Threats to Identification

The key identifying assumption is that no shock other than the fee reform differentially affected the September–August discontinuity in the treated cohort. Four threats warrant discussion. First, concurrent policy: the contemporaneous exemption of AAB+ students from number controls, the

raising of the participation age, and changes to maintenance support are set out in Section 3.3, and are addressed respectively by sub-AAB heterogeneity analysis, the choice of estimation window, and the design itself, which differences out changes that evolve smoothly across cohorts.

Second, strategic delay of school entry. Some summer-born children defer school entry to avoid the relative age effect, a practice concentrated among higher-income families. For this to bias the results, there would need to be a discontinuous, cohort-specific change in delayed school entry for children who would later face the 2012 fees. That is, among children starting school around 1998. Figure A.3 plots the share of students delaying primary school entry by university-entry cohort: the proportion is stable at approximately 0.5 percent, with no discontinuity around the treated cohort. As an additional precaution, the main analysis excludes individuals whose school-entry timing does not align with their birthdate cohort.

Third, seasonality of birth. Family background varies with season of birth (Buckles and Hungerman 2013), and the RAE premium itself partly reflects this selection. The design does not require births to be random across months, only that the selection pattern is stable across cohorts, in which case it is differenced out. The balance results above and the stability of the placebo discontinuities are direct evidence of this stability.

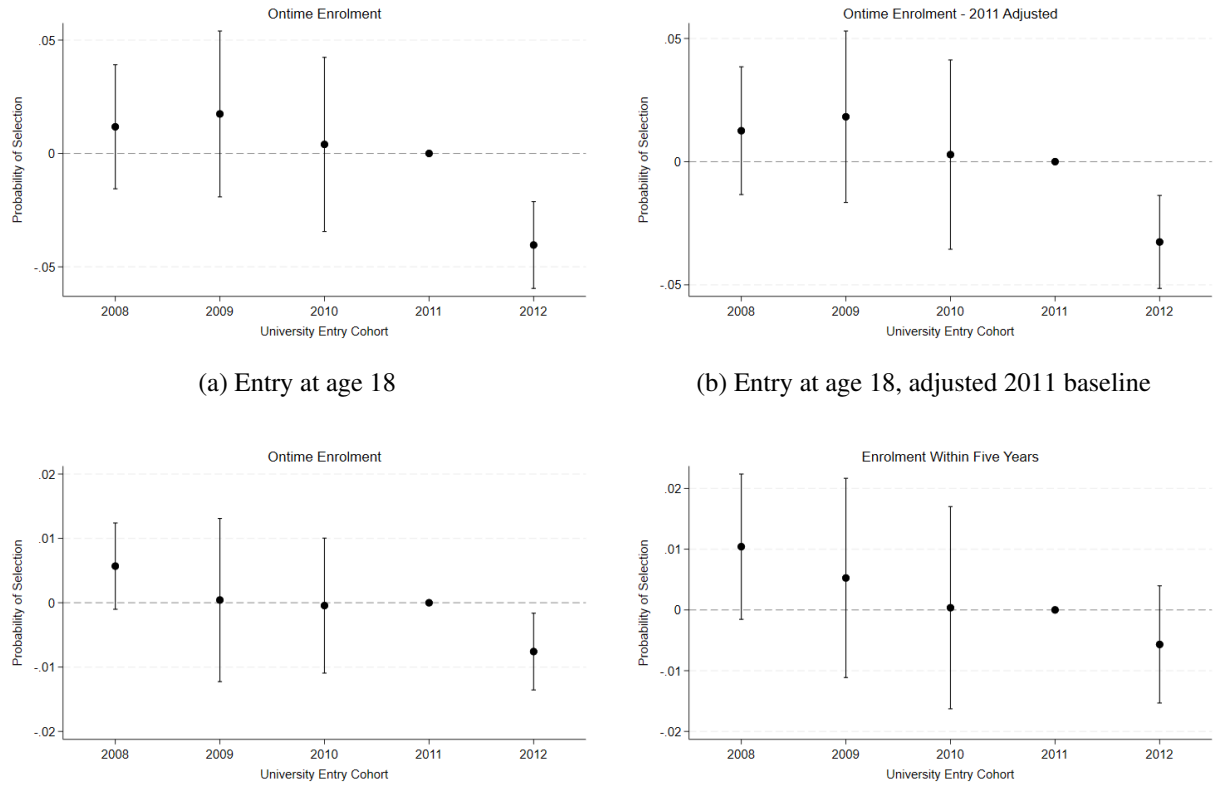
6 Results

6.1 Enrolment: Retiming, Not Deterrence

Figure 4 presents the paper's central result. Panel (a) plots the estimated difference-in-discontinuity effects on entry at age 18 for each cohort. The placebo estimates are tight around zero, indicating a stable relative age effect in the pre-reform period, while the treated cohort exhibits a sharp decline of around 4 percentage points — a 16% fall in on-time entry, and well above the 2.3 percentage point raw decline between the 2011 and 2012 entry classes. Following the estimand definitions in Section 5.1, this is the *total* effect of the reform and its announcement, and conflates

the permanent deterrence of entry with retiming induced by the two-year announcement window. The remaining panels decompose it.

Figure 4



(a) Entry at age 18

(b) Entry at age 18, adjusted 2011 baseline

(c) Entry by age 21

(d) Entry by age 23

Note: Figure plots the estimated difference-in-discontinuity effects on university entry by cohort. Panel (a) shows entry in the academic year following the individual's 18th birthday. Panel (b) re-estimates panel (a) against a counterfactual 2011 entry rate predicted from pre-announcement cohorts using a logit of entry on observable characteristics and a linear cohort trend. Panels (c) and (d) show cumulative entry within three and five years respectively, measures that are unaffected by retiming by construction. In panel (d), estimates for cohorts whose entry windows partially overlap the reform are shown in grey; see text. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

The announcement response is directly visible in the data: Figure A.1 in the appendix plots the share of entrants deferring entry by cohort: deferral collapses by 2 percentage points in 2011, as the final pre-reform cohort abandoned gap years to lock in the lower fees, exactly as the incentives described in Section 3 predict. A naïve correction follows by accounting: if the control cohort's entry at 18 was inflated by up to 2 percentage points of forgone deferral, the

treated cohort's own on-time response is bounded below at approximately 2 percentage points. Even this most conservative reading leaves a substantial contraction in on-time entry.

That reading is, however, almost certainly too conservative, because it assumes the pulled-forward deferrers raised the 2011 entry level one-for-one. They could not have: student number controls were binding in 2011, a year of record applications and record rejections, so additional demand from deferrers holding confirmed places displaced marginal applicants rather than expanding entry. Consistent with displacement, the 2011 entry class shows no excess in level above its trend. Panel (b) formalises the point: predicting counterfactual 2011 entry from pre-announcement cohorts using observable characteristics and a linear cohort trend[3], the predicted rate lies *above* the observed one, and the resulting estimate of the total effect rises to 4.4 percentage points. The bracketing is therefore [2.0, 4.4] percentage points, with the evidence favouring the upper region: the deferral collapse distorted who entered in 2011, not how many, leaving the unadjusted 4 percentage point estimate approximately correct as the total effect.

Panels (c) and (d) then ask how much of this contraction was permanent. Cumulative entry measures are immune to retiming by construction. A student entering at 19 rather than 18 counts identically on either side of the cutoff. So any remaining effect reflects students who did not enter at all within the horizon. Panel (c) shows entry within three years: the placebo estimates are flat and tight around zero, and the treated cohort declines by [X] percentage points, an estimate that lies outside the placebo distribution and survives at the 1% level. The permanent effect is thus a fraction of the total: of the roughly 4 percentage point contraction in on-time entry, approximately [X] percentage points, [X] percent of the response, reflects delayed entry rather than deterred entry.

Panel (d) extends the horizon to five years and requires more care. While the point estimate remains the same as in panel (c), there appears to be a downward-sloping pattern at this horizon. It is unclear what causes this pattern; it could be evidence of the relative age effect declining

[3]Gender, free school mean status, school level geography, a September birth indicator, and a linear cohort time trend are used as covariates in this regression

over time, it could be because each cohort becomes progressively treated as they age into the post-reform period, or it could be a statistical artifact. Fortunately, the retiming interpretation receives independent confirmation from outside the enrolment data entirely. If treated students shifted entry to later ages, the cohort should exhibit higher employment at ages 19–21, when delayed entrants are working before enrolment, and lower employment at ages 23–24, when they remain in study while their counterfactual on-time peers have graduated. Section 6.5 shows the employment age profile displays exactly this reversal, with magnitudes consistent with the entry-timing shifts estimated here.

Given the potential statistical caveats to the result in panel (d), the descriptive evidence presented in Section 3.2, and the broader literature, I would regard the 'null' result (as established by the inference used within this paper) as a lower bound for the permanent deterrence effect rather than the absence of an effect.

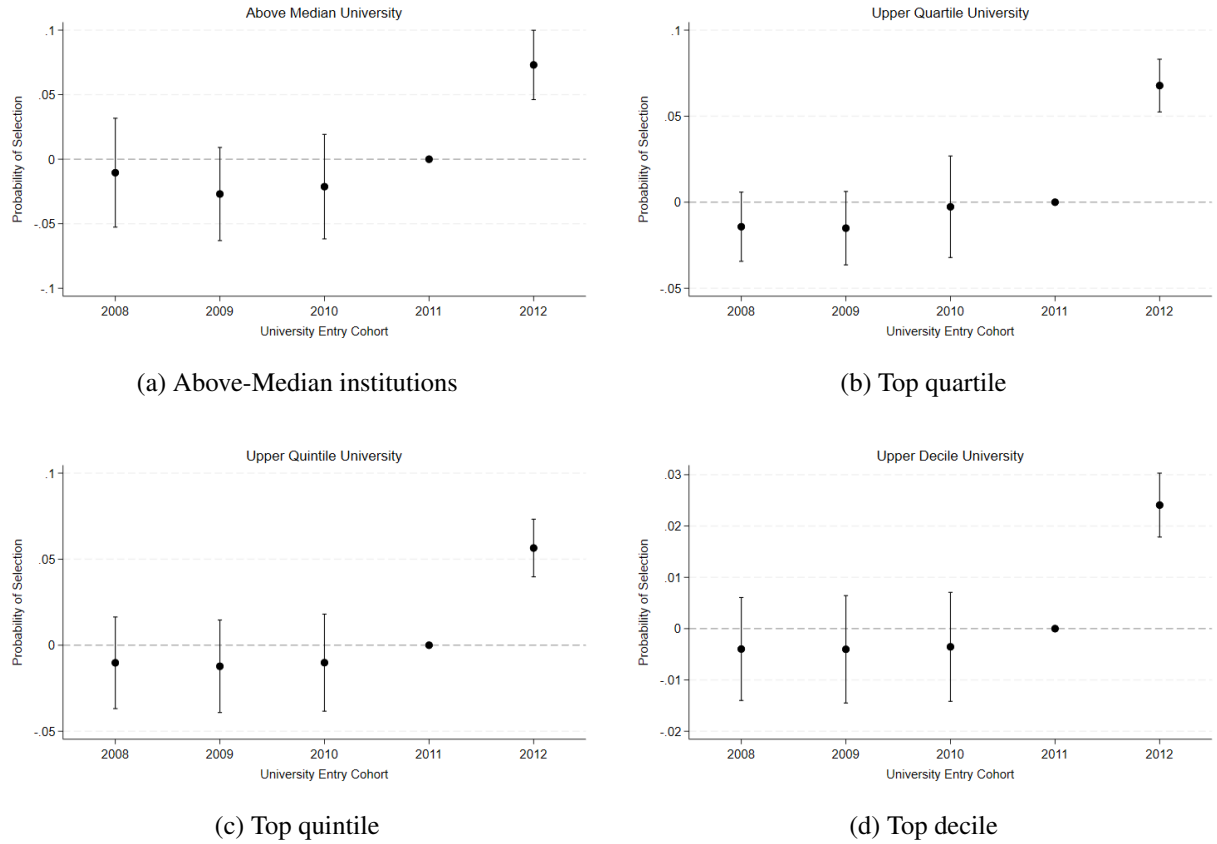
6.2 Reallocation Across Institutions

Having established that most of the enrolment response was a change in timing, I turn to where students enrolled. Figure 5 plots difference-in-discontinuity estimates for entry into more selective institutions, conditional on university entry, across four increasingly demanding tiers: institutions above the mean of the selectivity distribution, and in its top quartile, quintile, and decile.[4] All outcomes are ever-entered measures and therefore, like the cumulative entry results above, immune to retiming by construction. The placebo estimates are tight around zero in every panel, while the treated cohort shows a significant increase at every tier: conditional on attending, students were [X] percentage points more likely to enter an above-mean institution, [X] points more likely to enter the top quartile, [X] the top quintile, and [X] the top decile. The effect is not confined to the margin of the median institution but extends to the most selective end of the

[4]Selectivity is defined by ranking institutions according to the average A-level points of their entering students; Section 7 shows the results are robust to the choice of reference year.

distribution.

Figure 5



(a) Above-Median institutions

(b) Top quartile

(c) Top quintile

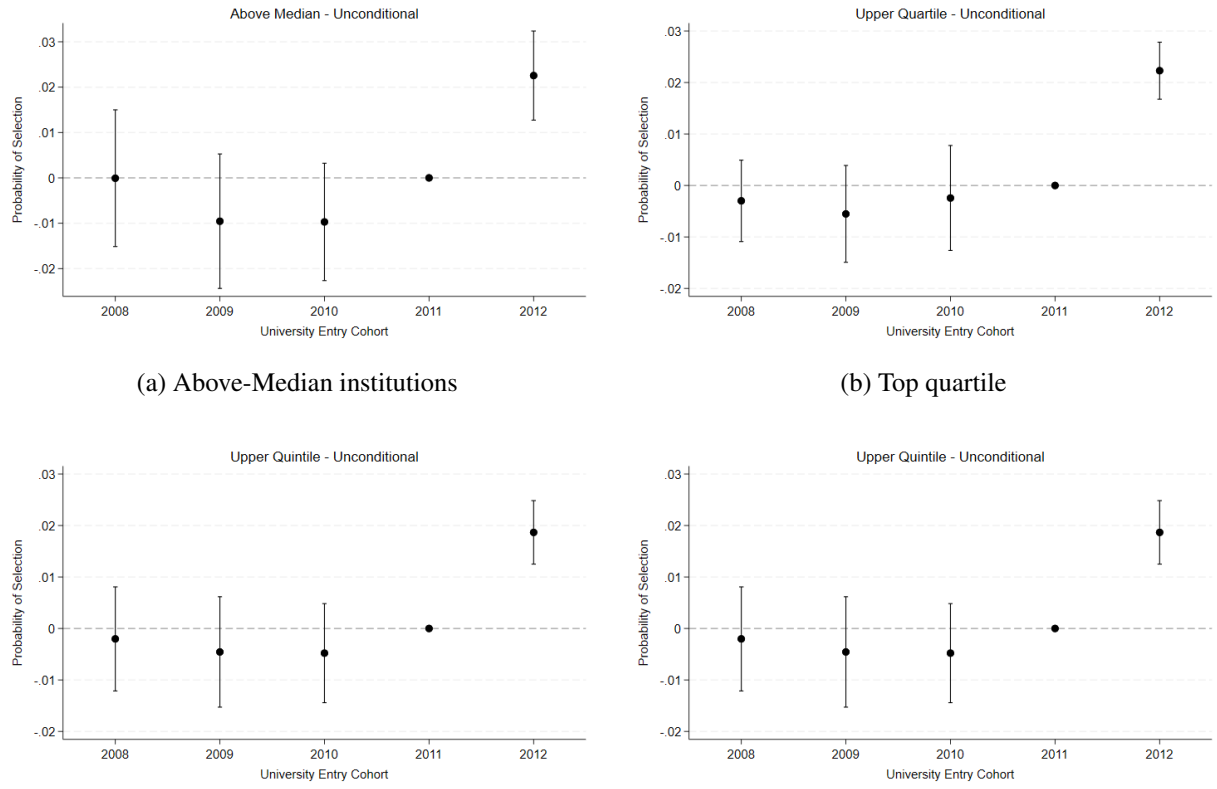
(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on ever-entry into institutions in each selectivity tier, conditional on university entry. Selectivity tiers are defined by the average A-level points of entering students. Ever-entered measures are unaffected by retiming by construction. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

The conditional estimates alone cannot distinguish genuine upgrading from a compositional artefact: if the fee increase deterred marginal students who would have attended less selective institutions, the remaining entry class would look more selective even with no individual changing destination. Figure 6 therefore re-estimates each tier unconditionally, coding the outcome as zero for individuals who never enter university, so that the estimates measure the probability that a member of the cohort ever attends a selective institution. The pattern is unchanged: smaller in magnitude, as expected, but positive and significant at every tier, with unconditional increases of [X], [X], [X], and [X] percentage points respectively. Selection into higher educa-

tion became slightly less common in the treated cohort; entry into its selective segment became unambiguously more common.

Figure 6



(a) Above-Median institutions

(b) Top quartile

(c) Top quintile

(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on conditional ever-entry into each selectivity tier, coded zero for individuals who never enter university. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

Combining the unconditional estimates with the enrolment results yields the paper's central piece of arithmetic. Total ever-entry by age 23 fell by only [X] percentage points, while unconditional ever-entry into above-mean institutions *rose* by [X] points; entry into below-mean institutions therefore fell by approximately [X] points — the sum of the two. The fee increase operated principally as a reallocation: a cohort only marginally smaller in the end attended a markedly different set of institutions, shifting up the quality distribution, with the contraction concentrated entirely in its less selective segment. Notably, the unconditional increase at the top

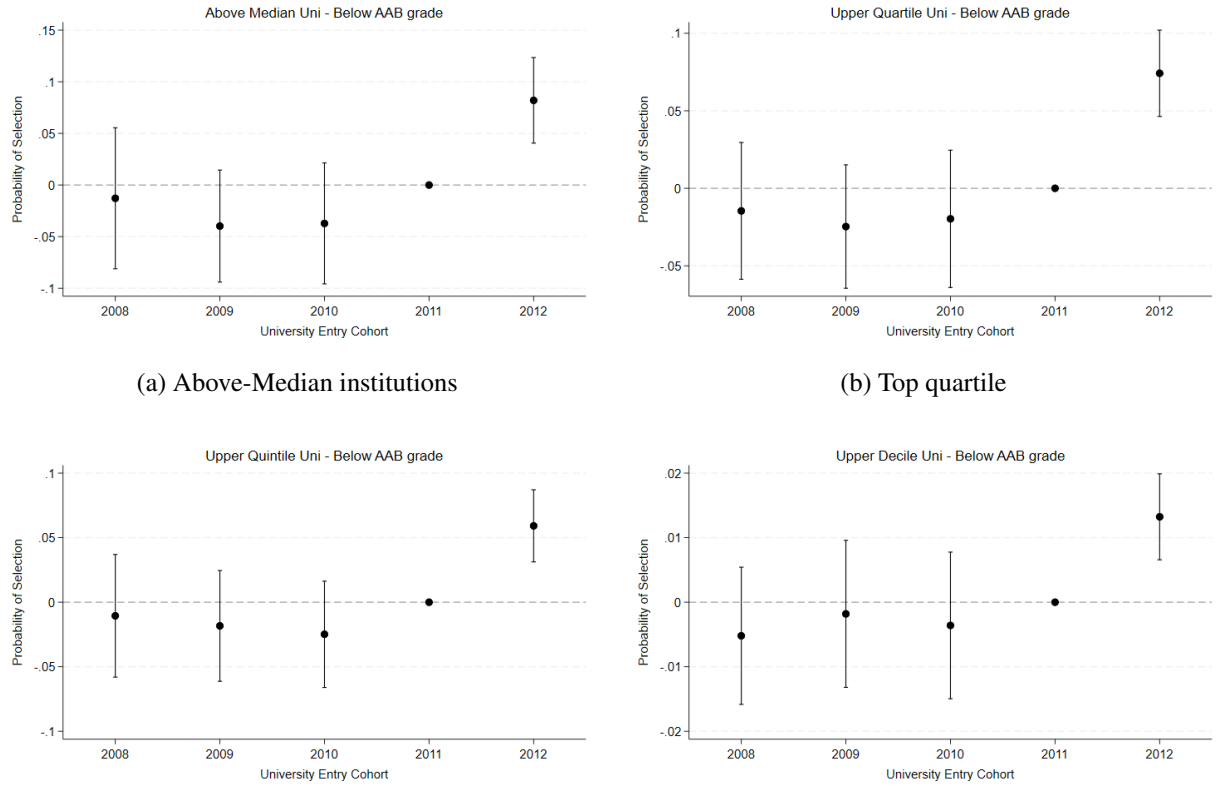
implies that more selective institutions expanded the number of places taken by the cohort in absolute terms, consistent with the reform operating as an increase in a binding price control on institutions whose willingness to supply places at the regulated price exceeded their allocations.

Two threats to this interpretation must be ruled out, one on each side of the cutoff. The first concerns the control cohort. Students pulling entry forward into 2011 to avoid the higher fees faced a binding number cap and record competition for places (Section 6.1); some may have accepted offers from less selective institutions than they would otherwise have attended, simply to secure entry under the old regime. Such anticipatory downgrading would depress the control cohort's selective entry rate and mechanically generate the treated-cohort increase. Figure A.2 in the appendix shows directly that no such dip occurred: entry into above-mean institutions in 2011 sits on its pre-announcement trend, with no decline in the announcement years. As further reassurance, Figures A.4 and **??** re-estimate both sets of tier results against the adjusted 2011 baseline constructed in Section 6.1; the estimates are essentially unchanged. If anything, the displacement documented earlier, which strengthened the composition of the 2011 entry class, raises the control-side bar for selective entry and renders the estimates here conservative.

The second threat concerns the treated cohort, and is the concurrent policy described in Section 3.3: from 2012–13, students achieving AAB or higher at A-level were exempted from student number controls, freeing selective institutions to expand recruitment of high-attaining students at precisely the moment the fee reform took effect. If the reallocation documented above merely reflects this supply-side release, it should be concentrated among AAB+ students. Figure A.5 re-estimates the unconditional tier results for students with attainment *below* the AAB threshold, who were unaffected by the exemption. Figure A.6 shows the same pattern for the unconditional enrolment of student who achieved less than AAB. The increase in selective entry remains positive and significant at each tier, demonstrating that the reallocation operated through the fee reform itself rather than through the exemption. The natural interpretation ties the two results together: the fee increase functioned as a rise in a binding price cap, raising the regulated price

toward institutions' willingness to supply places, and selective institutions expanded along margins well below the exemption threshold.

Figure 7



(a) Above-Median institutions

(b) Top quartile

(c) Top quintile

(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on unconditional ever-entry into each selectivity tier for students with A-level attainment below AAB, who were unaffected by the 2012–13 exemption of AAB+ students from student number controls. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

6.3 Attainment

If the reform reshuffled who entered university, when, and where, a natural concern is that the reshuffled students performed worse: that delayed entrants lost momentum, or that students trading up to more selective institutions found themselves academically overmatched. Figure 8 examines four attainment outcomes among entrants: completion of the degree, outright failure, enrolment in a STEM subject, and progression to postgraduate study.

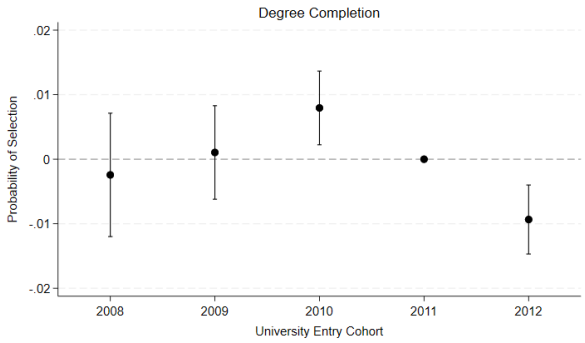
Panels (a) and (b) show degree completion and failure. In both cases the placebo estimates drift gently upward across pre-reform cohorts, consistent with the secular improvement in completion rates over this period, and in the treated cohort the coefficient turns negative. Neither estimate, however, is significant at the 1 percent threshold adopted throughout, and neither lies outside the placebo distribution: by the paper's own inferential standards these are null results. The point estimates bound the scope for harm at [X] percentage points for completion and [X] for failure, magnitudes small relative to baseline rates of [X] and [X] percent. The students who delayed entry or entered more selective institutions did not do detectably worse, a finding consistent with the reallocation reflecting improved matching rather than overreach.

Panels (c) and (d) report subject choice and postgraduate progression, and I place little weight on either. The STEM estimates are imprecise across all cohorts, and the postgraduate estimates appear to load on the strong secular growth in postgraduate enrolment over the period, including the expansion of master's provision and the 2016 introduction of postgraduate loans, which the design cannot cleanly difference out at the end of the sample window. Both panels are reported for completeness: they show no evidence that the reform shifted students toward or away from STEM subjects or further study, but the estimator is poorly suited to these outcomes and the results should be read as uninformative rather than as precise nulls.

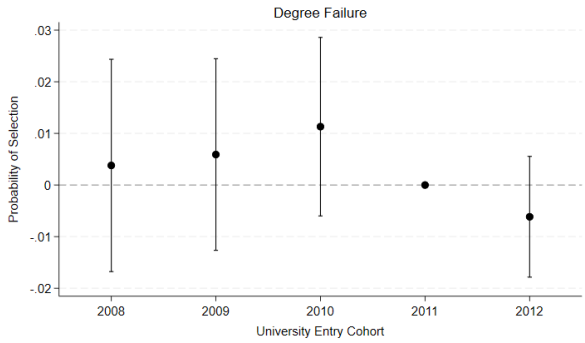
6.4 Heterogeneity

The reallocation documented above could conceal uneven incidence: if price sensitivity varies with family income, gender, neighbourhood, or prior attainment, an aggregate reallocation could coexist with widening participation gaps for particular groups. The progressive design of the loan system predicts otherwise. Because repayment is contingent on income, the expected cost of the fee increase is lowest for students with the lowest expected earnings, and lowest of all for those who expect never to cross the repayment threshold; if anything, the reform should bind least on the most disadvantaged.

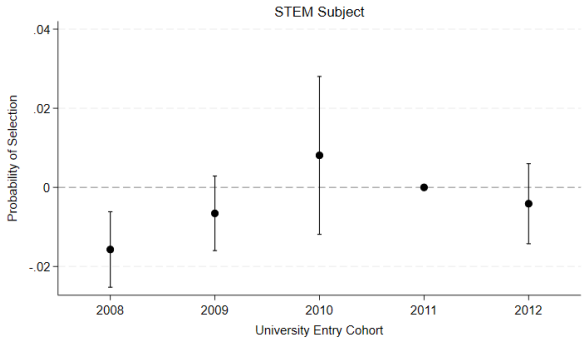
Figure 8



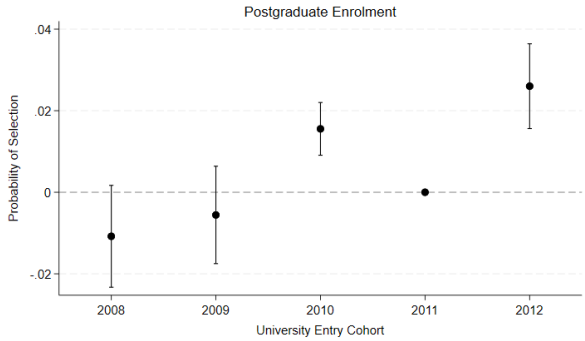
(a) Degree completion



(b) Degree failure



(c) STEM enrolment



(d) Postgraduate progression

Note: Figure plots the estimated difference-in-discontinuity effects on attainment outcomes among university entrants. Panels (a) and (b) show completion of the degree and outright failure. Panels (c) and (d) show enrolment in STEM subjects and progression to any postgraduate course; as discussed in the text, these estimates are imprecise and, in the case of postgraduate study, likely confounded by secular growth in provision at the end of the sample window. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

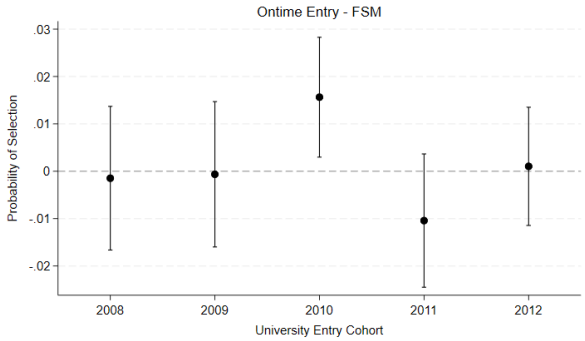
Figure 9 tests this by interacting the difference-in-discontinuities estimator with four characteristics: free school meal eligibility, gender, residence in a low university participation area, and standardised KS5 attainment. No interaction is significant at the 1 percent threshold in the treated cohort, and no treated-cohort interaction lies outside its own placebo distribution. The confidence intervals rule out differential responses larger than [X] percentage points for FSM students, [X] for women, [X] by neighbourhood participation, and [X] per standard deviation of attainment. The enrolment response was, as far as the data can distinguish, uniform across these dimensions.

Two readings of this uniformity deserve emphasis. First, on the equity question that dominates the policy debate, there is no evidence that the tripling of fees widened participation gaps by income, gender, or neighbourhood, consistent with the income-contingent loan system insulating disadvantaged students from a price they would only ever repay out of graduate earnings. Second, the absence of an attainment gradient is itself informative about mechanism. A pure demand-side story in which fees price out the marginal student would predict the response concentrating among lower attainers; its absence is instead consistent with the reallocation operating substantially through the supply side, as institutions responded to the higher price cap, a reading that matches the sub-AAB evidence in Section 6.2. I do not push this interpretation further than the precision of the interactions allows, but the two findings point the same way.

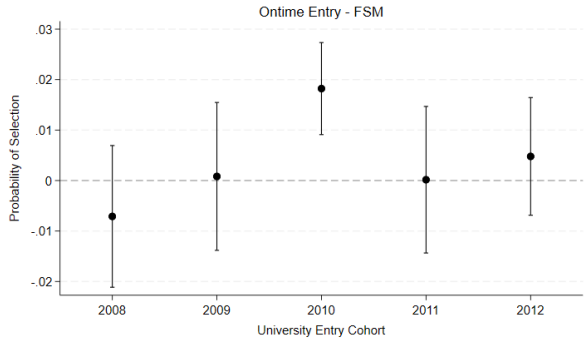
6.5 Labour Market Dynamics

The retiming interpretation makes a sharp out-of-sample prediction. If the treated cohort shifted university entry to later ages without materially reducing eventual attendance, its members should be observed working more in the years when delayed entrants had not yet enrolled, and working less in the years when they remained in study while their counterfactual on-time peers had graduated, with the differences fading once the delayed entrants themselves graduate. The prediction is testable in data that played no role in the enrolment analysis: the HMRC employment and earnings records.

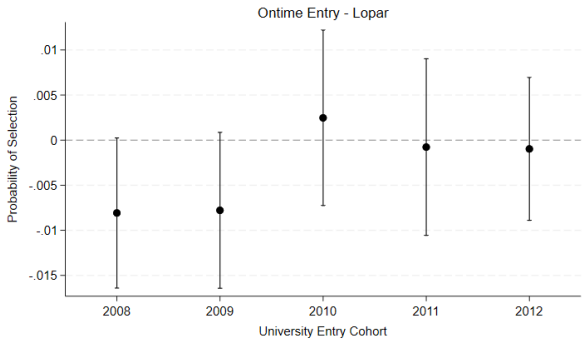
Figure 9



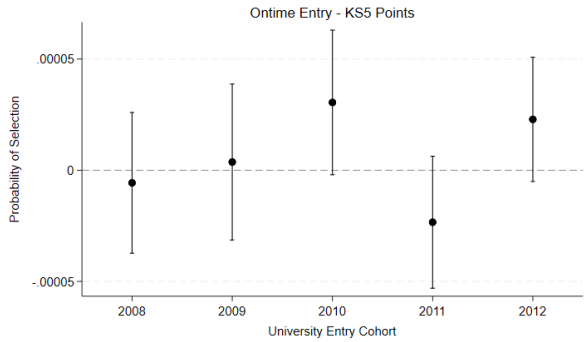
(a) Free school meal eligibility



(b) Gender



(c) Low university participation area



(d) KS5 attainment

Note: Figure plots the estimated difference-in-discontinuity effects on university entry interacted with each characteristic. Each panel reports the interaction coefficient by cohort: a significant treated-cohort estimate would indicate a differential response for the named group. FSM = Free School Meals. KS5 attainment is standardised total points at A-level. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

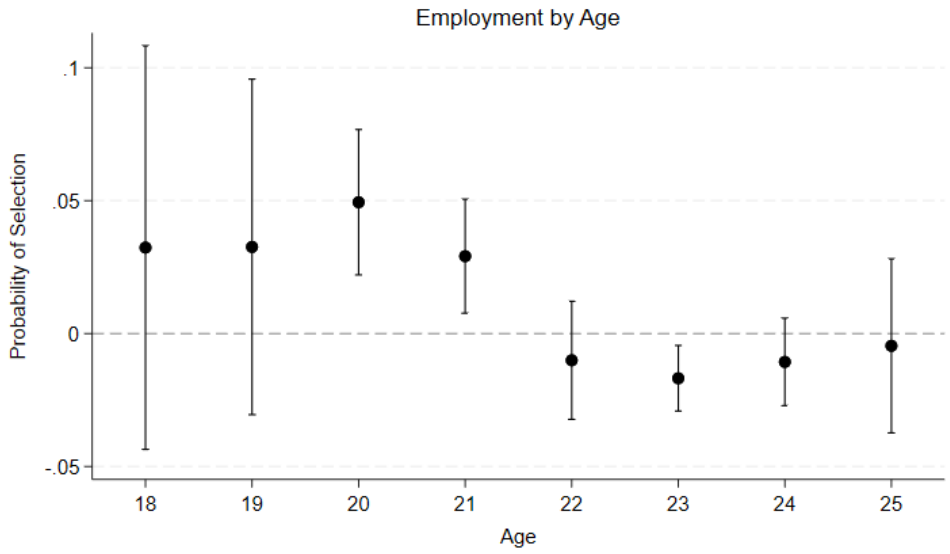
Figures 10 and 11 plot difference-in-discontinuity estimates on employment and earnings at each age from 18 to 25. Both display the same reversal. Employment and earnings are higher for the treated cohort at ages 19 and 20, when a larger share of the cohort was in work rather than study; the differences die out at 21, turn negative at ages 22 and 23, when delayed entrants remained enrolled while their on-time counterparts entered the labour market; and fade to zero by ages 24 and 25, as the delayed entrants complete their degrees and catch up. The profile is the mirror image of the entry-timing shifts estimated in Section 6.1, recovered from tax records rather than education records. I treat this as independent confirmation of the retiming mechanism: no plausible confound operating through the labour market alone would generate a reversal timed precisely to the cohort's delayed progression through university.

The endpoint of the profile carries its own substantive weight. By ages 24 and 25 the treated cohort's employment and earnings are statistically indistinguishable from the placebo cohorts', with estimates tight around zero. Whatever disruption the reform caused to the timing of entry, and whatever reallocation it induced across institutions, there is no evidence of scarring in early career outcomes: the cohort that paid nearly triple the fees arrived at its mid-twenties working and earning at the same rates as its predecessors. A fuller assessment of earnings effects, in particular any returns to the reallocation toward more selective institutions documented in Section 6.2, requires observing the cohort deeper into its career, since graduate earnings premia widen with age and the most selective institutions' premia widen most. I leave this to future work as the data horizon extends.

Two measurement concerns are addressed in the appendix. First, the oldest ages in the profile are observed in tax years affected by the COVID-19 pandemic, which disrupted employment for young workers broadly. Figures A.7 and A.9 re-estimate the profiles excluding tax years from 2019–20 onwards; the reversal pattern at ages 19 to 23 is unchanged, and the convergence at the oldest ages is unaffected where it remains observable. The earnings measure is conditional on employment, which makes the two profiles complementary rather than duplicative. The em-

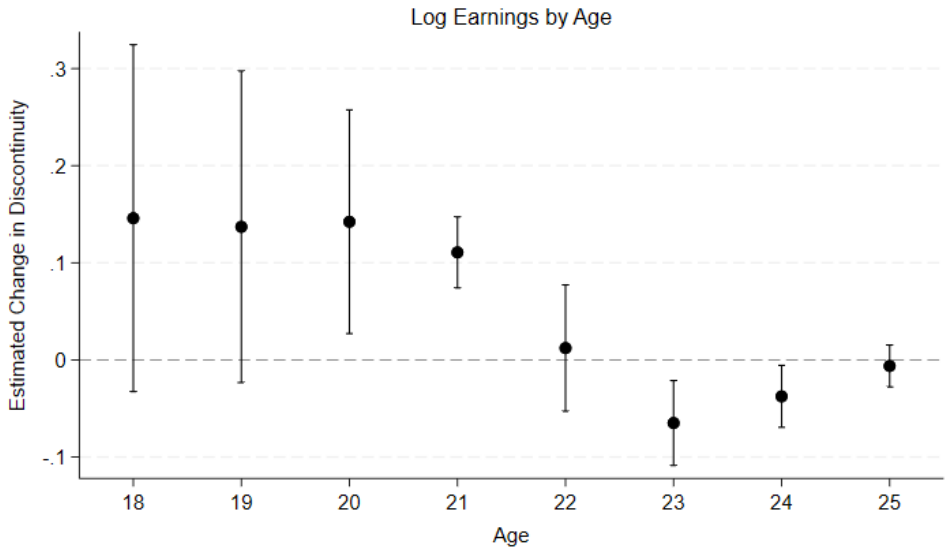
ployment profile captures the extensive margin: more of the treated cohort in work at 19 and 20, fewer at 22 and 23. The earnings profile adds the intensive margin: conditional on working, treated individuals earned more at ages 19 and 20, consistent with delayed entrants holding full-time employment before enrolment rather than the part-time work typical of students, and earned less at ages 22 and 23, when the treated cohort's employed pool contained more still-enrolled students in part-time work while their on-time counterparts had entered graduate employment. Composition of the employed pool therefore differs across the cutoff by construction; this is precisely the mechanism the profile is designed to detect, but it means the convergence in conditional earnings at ages 24 and 25 should be read jointly with the employment convergence rather than as an independent result. Together the two margins tell one story: the treated cohort worked more, and more fully, before its delayed entry, and less, and less fully, while completing it.

Figure 10



Note: Figure plots the estimated difference-in-discontinuity effects on employment at each age from 19 to 25. Positive estimates at ages 19 and 20 reflect delayed university entrants in work before enrolment; negative estimates at ages 22 and 23 reflect delayed entrants remaining in study while on-time peers entered the labour market; estimates converge to zero by ages 24 and 25. Data come from the 1998–2020 National Pupil Database, the Higher Education Statistical Agency census, and HMRC employment records.

Figure 11



Note: Figure plots the estimated difference-in-discontinuity effects on earnings at each age from 19 to 25, including zeros for individuals not in employment. The profile mirrors the employment profile in Figure 10 and matches the entry-timing shifts estimated in Section 6.1. Data come from the 1998–2020 National Pupil Database, the Higher Education Statistical Agency census, and HMRC earnings records.

7 Robustness

The paper's inferential architecture places the burden of proof on the placebo distribution, so the robustness of the results turns on two questions: whether the placebo estimates remain tight around zero, and the treated estimates far outside them, as the researcher degrees of freedom in the design are varied; and whether the checks embedded in the results sections survive scrutiny. This section addresses the main specification choices in turn. Throughout, a result is treated as robust only if the treated estimate remains outside the span of the placebo estimates under the alternative specification, the same standard applied to the central results.

The first specification choice is the bandwidth around the 1 September cutoff. The central estimates compare individuals born in September with those born in August, the narrowest window and the cleanest comparison in age. Widening the window trades this sharpness for statistical precision, at the cost of comparing individuals further apart in age and further from the cutoff. I re-estimate all central results defining the discontinuity over two and three months on either

side of the cutoff. The enrolment and reallocation results are highly stable: the placebo estimates remain centred on zero, and the treated estimates remain of similar magnitude and outside the placebo span at every bandwidth. The labour market age profiles retain their shape, with the reversal at ages 19 to 23 visible at all bandwidths, though the estimates are noisier as the comparison groups become less sharply defined in age. No conclusion in the paper depends on the choice of a one-month window.

The second choice is the reference year used to rank institutions by selectivity. The central results rank institutions by the average A-level points of their entering students, and a natural concern is that the ranking is itself an outcome: if the reform changed which students attended which institutions, a ranking constructed from post-reform entry classes would embed the treatment. I therefore fix the ranking using pre-reform entry cohorts and show that the tier results are insensitive to the choice of base year, re-estimating the reallocation results with rankings constructed from each available pre-reform year. The estimates are essentially unchanged, reflecting the stability of the institutional hierarchy itself.

Third, the checks embedded in the results sections are collected here for reference. The contamination of long-horizon cumulative outcomes by progressively treated controls is diagnosed, predicted, and shown in the graph of ever-entry (by age 23), leading to a cautious interpretation of ever-enrolment as a lower bound. However, I believe the labour market results independently showing traces of the retiming provide robust evidence of the retiming mechanism. The concurrent AAB exemption is ruled out as the source of the reallocation by the sub-AAB estimates (Section 6.2). The share of children with delayed school entry is stable at approximately 0.5 percent across all cohorts with no discontinuity around the treated cohort (Figure A.3), ruling out compositional and strategic sorting explanations.

Finally, the placebo framework itself deserves a candid statement of its limits. With seven placebo cohorts, the design can demonstrate that the treated estimates are unlike anything the placebo process produces, but it cannot attach conventional p-values to that statement, and results

whose treated estimates sit near the edge of the placebo span, such as the attainment coefficients in Section 6.3, are correctly reported as nulls rather than marginal findings. The discipline cuts in both directions: it is what prevents overclaiming on the weaker results, and what makes the central results credible, since the enrolment, reallocation, and labour market estimates exceed the largest placebo in their respective distributions.

8 Conclusion

This paper asks what happened when the price of university in England nearly tripled. Exploiting the alignment of the 2012 fee reform with the September school-entry cutoff, and drawing on population-wide administrative data that follow entire cohorts from school through university entry at any age and into the labour market, I decompose the enrolment response to a large, pre-announced price increase into its temporary and permanent components, and trace where the affected cohort ultimately enrolled.

The answer is that the reform changed the timing and destination of university entry far more than its level. Entry at age 18 fell by around four percentage points, but most of this contraction reflects retiming: the final pre-reform cohort abandoned deferral to lock in the lower fees, and treated students, facing the higher fees whenever they entered, delayed. Cumulative entry within three years fell by only [X] percentage points, an estimate stable across horizons once the progressive treatment of control cohorts' late entrants is accounted for. The labour market independently confirms the mechanism: the treated cohort worked more, and in fuller employment, at ages 19 and 20, worked less at ages 22 and 23 while its delayed entrants remained in study, and was indistinguishable from its predecessors by ages 24 and 25. The cohort that faced £9,000 fees was largely delayed, and modestly deterred.

What the reform did change permanently is where students enrolled. Ever-entry into more selective institutions rose across the entire selectivity distribution, both conditional on atten-

dance and unconditionally, implying that a cohort only marginally smaller in the end attended a markedly different set of institutions, with the contraction concentrated entirely in the less selective segment. This reallocation is not explained by the concurrent exemption of high-attaining students from number controls, since it holds among students below the AAB threshold, and not by anticipatory downgrading in the control cohort, whose selective entry sits on trend. The natural reading is that the reform operated as an increase in a binding price control: selective institutions, willing to supply more places at the higher regulated price, expanded, and students moved up the quality distribution to fill them. Alongside these shifts, I find no detectable effects on degree completion or failure, no heterogeneity in the enrolment response by family income, gender, neighbourhood participation, or prior attainment, and no scarring in employment or earnings by the mid-twenties.

The policy implications are specific to the institutional setting, and the setting is the point. Every effect documented here operated through a system in which no student paid fees upfront and repayment was contingent on income. Under those conditions, a near-tripling of the headline price did not measurably widen participation gaps, did not durably reduce participation by more than [X] percentage points, and shifted enrolment toward more selective institutions. This does not license the conclusion that fees are costless: the reform transferred a substantial share of higher education funding from taxpayers to graduates, with distributional consequences across the earnings distribution of graduates themselves that this paper does not evaluate, and the deterred margin, though small, is real. What the evidence does support is narrower and still consequential: in the presence of universal income-contingent loans, the participation margin is far less sensitive to the fee level than the public debate assumes, and the margins that do respond, timing and institutional destination, are largely absent from that debate.

The findings also locate the UK model in international context, with more precision than advocacy on either side usually allows. Systems that rely on upfront prices and private credit expose the participation margin directly, and the large literature on credit constraints shows that margin

responds. Systems with no fees protect participation but forgo the price signal entirely. The English arrangement, high fees deferred through income-contingent loans, appears in this episode to have suppressed the participation response while leaving the allocation response active. Whether that allocation response improved welfare depends on the returns to the institutional reallocation it induced, which cannot be evaluated at age 25 and which I leave to future work as the data horizon extends.

The limitations are worth stating plainly. The design identifies the effect of one large reform in one system, from a single treated cohort, with inference resting on the stability of placebo discontinuities rather than conventional standard errors; the permanent enrolment effect, though robust to the contamination corrections developed above, is bounded rather than point-identified with precision; and the labour market horizon ends in the cohort's mid-twenties, before the returns to institutional selectivity fully materialise. Within those limits, the paper offers the first decomposition of a national tuition fee reform into its retiming and deterrence components, and the first evidence that such a reform reallocated a cohort across the quality distribution of institutions.

The broader lesson for the economics of higher education finance is about margins. The canonical question, does price reduce participation, receives a muted answer in this setting precisely because the loan system was built to mute it. The active margins were the ones a participation regression cannot see: when students enrolled, and where. Reforms to higher education finance should be evaluated, and designed, with those margins in view.

References

Azmat, Ghazala and Ştefania Simion, "Charging for higher education: Estimating the impact on inequality and student outcomes," *The BE journal of economic analysis & policy*, 2021, *21* (1), 175–239.

Becker, Gary S, "Human capita," *New York: N ational Bureau of Economic R esearch*, 1964.

Bettinger, Eric, "How financial aid affects persistence," in "College choices: The economics of where to go, when to go, and how to pay for it," University of Chicago Press, 2004, pp. 207–238.

Black, Sandra E, Jeffrey T Denning, Lisa J Dettling, Sarena Goodman, and Lesley J Turner, "Taking it to the limit: Effects of increased student loan availability on attainment, earnings, and financial well-being," *American Economic Review*, 2023, *113* (12), 3357–3400.

_ , Paul J Devereux, and Kjell G Salvanes, "Too young to leave the nest? The effects of school starting age," *The review of economics and statistics*, 2011, *93* (2), 455–467.

Britton, Jack, Lorraine Dearden, and Ben Waltmann, "The 2012 tuition fee reforms made the poorest graduates £1,500 better off. Later reforms have more than wiped out that gain," https://ifs.org.uk/news/2012-tuition-fee-reforms-made-poorest-graduates-ps1500-better-reforms-have-more-wiped-out-gain 2017. Accessed July 2025.

_ , _ , Laura van der Erve, and Ben Waltmann, "The impact of undergraduate degrees on lifetime earnings," Technical Report, IFS Report 2020.

Bruckmeier, Kerstin and Berthold U Wigger, "The effects of tuition fees on transition from high school to university in Germany," *Economics of education review*, 2014, *41*, 14–23.

Buckles, Kasey S and Daniel M Hungerman, "Season of birth and later outcomes: Old questions, new answers," *Review of economics and statistics*, 2013, *95* (3), 711–724.

Callender, Claire and Jonathan Jackson, "Does the fear of debt deter students from higher education?," *Journal of social policy*, 2005, *34* (4), 509–540.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Campbell, Tammy, "In-school ability grouping and the month of birth effect," *Preliminary evidence from the Millenium Cohort Study. London: Centre for Longitudinal Studies*, 2013.

Castleman, Benjamin L and Bridget Terry Long, "Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation," *Journal of Labor Economics*, 2016, *34* (4), 1023–1073.

Committee on Higher Education, "Higher Education: Report of the Committee appointed by the Prime Minister under the Chairmanship of Lord Robbins," Command Paper, London 1963. Cmnd. 2154.

Conley, Timothy G. and Christopher R. Taber, "Inference with "Difference in Differences" with a Small Number of Policy Changes," *The Review of Economics and Statistics*, 2011, *93* (1), 113–125.

Crawford, Claire, Lorraine Dearden, and Ellen Greaves, *When you are born matters: evidence for England* number R80, IFS Report, 2013.

_ , _ , and _ , "The drivers of month-of-birth differences in children's cognitive and non-cognitive skills," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 2014, *177* (4), 829–860.

Dearden, Lorraine, Emla Fitzsimons, and Gill Wyness, "The Impact of Tuition Fees and Support on University Participation in the UK. CEE DP 126.," *Centre for the Economics of Education (NJ1)*, 2011.

Deming, David and Susan Dynarski, "Into college, out of poverty? Policies to increase the postsecondary attainment of the poor," Technical Report, National Bureau of Economic Research 2009.

Denning, Jeffrey T, Benjamin M Marx, and Lesley J Turner, "ProPelled: The effects of grants on graduation, earnings, and welfare," *American Economic Journal: Applied Economics*, 2019, *11* (3), 193–224.

Dynarski, Susan M, "Does aid matter? Measuring the effect of student aid on college attendance and completion," *American Economic Review*, 2003, *93* (1), 279–288.

Field, Erica, "Educational debt burden and career choice: Evidence from a financial aid experiment at NYU Law School," *American Economic Journal: Applied Economics*, 2009, *1* (1), 1–21.

Hagemann, Andreas, "Placebo Inference on Treatment Effects When the Number of Clusters Is Small," *Journal of Econometrics*, 2019, *213* (1), 190–209.

Hemelt, Steven W and Dave E Marcotte, "The impact of tuition increases on enrollment at public colleges and universities," *Educational Evaluation and Policy Analysis*, 2011, *33* (4), 435–457.

Hillman, Nicholas, "From grants for all to loans for all: Undergraduate finance from the implementation of the Anderson Report (1962) to the implementation of the Browne Report (2012)," *Contemporary British History*, 2013, *27* (3), 249–270.

House of Commons Library, "Tuition Fee Statistics," Technical Report, House of Commons Library 2018. Notes that "most universities immediately began to charge the maximum £9,000 fee".

Hübner, Malte, "Do tuition fees affect enrollment behavior? Evidence from a 'natural experiment'in Germany," *Economics of Education Review*, 2012, *31* (6), 949–960.

Independent Review of Higher Education Funding and Student Finance, "Securing a Sustainable Future for Higher Education: An Independent Review of Higher Education Funding and Student Finance," Presented to Parliament by the Secretary of State, October 2010 2010. Chair: Lord Browne of Madingley.

Kane, Thomas J, "Rising public college tuition and college entry: How well do public subsidies promote access to college?," 1995.

Lochner, Lance J and Alexander Monge-Naranjo, "The nature of credit constraints and human capital," *American economic review*, 2011, *101* (6), 2487–2529.

MacKinnon, James G. and Matthew D. Webb, "The Wild Bootstrap for Few (Treated) Clusters," *The Econometrics Journal*, 2018, *21* (2), 114–135.

_ **and** _ , "Randomization Inference for Difference-in-Differences with Few Treated Clusters," *Journal of Econometrics*, 2020, *218* (2), 435–450.

Malani, Anup and Julian Reif, "Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform," *Journal of Public Economics*, 2015, *124*, 1–17.

McPhillips, Martin and Julie-Anne Jordan-Black, "The effect of month of birth on the attainments of primary and secondary school pupils," *British Journal of Educational Psychology*, 2009, *79* (3), 419–438.

Mincer, Jacob A, "Schooling and earnings," in "Schooling, experience, and earnings," NBER, 1974, pp. 41–63.

Murphy, Richard, Judith Scott-Clayton, and Gill Wyness, "The end of free college in England: Implications for enrolments, equity, and quality," *Economics of Education Review*, 2019, *71*, 7–22.

National Committee of Inquiry into Higher Education, "Higher Education in the Learning Society: Report of the National Committee of Inquiry into Higher Education," Report to the Secretary of State for Education and Employment, London 1997. Chair: Sir Ron Dearing.

Neill, Christine, "Tuition fees and the demand for university places," *Economics of Education Review*, 2009, *28* (5), 561–570.

Russell, RJH and MJ Startup, "Month of birth and academic achievement," *Personality and individual differences*, 1986, *7* (6), 839–846.

Sá, Filipa, "The effect of tuition fees on university applications: evidence from the UK," 2014.

UK-Government, "Teaching and Higher Education Act," Technical Report 1998.

UK Parliament, "Education Act 1962," https://www.legislation.gov.uk/ukpga/Eliz2/10-11/12/enacted 1962. 10 & 11 Eliz. 2 c. 12.

_ , "Higher Education Act 2004," https://www.legislation.gov.uk/ukpga/2004/8/enacted 2004. 2004 c. 8.

Appendix Figures and Tables

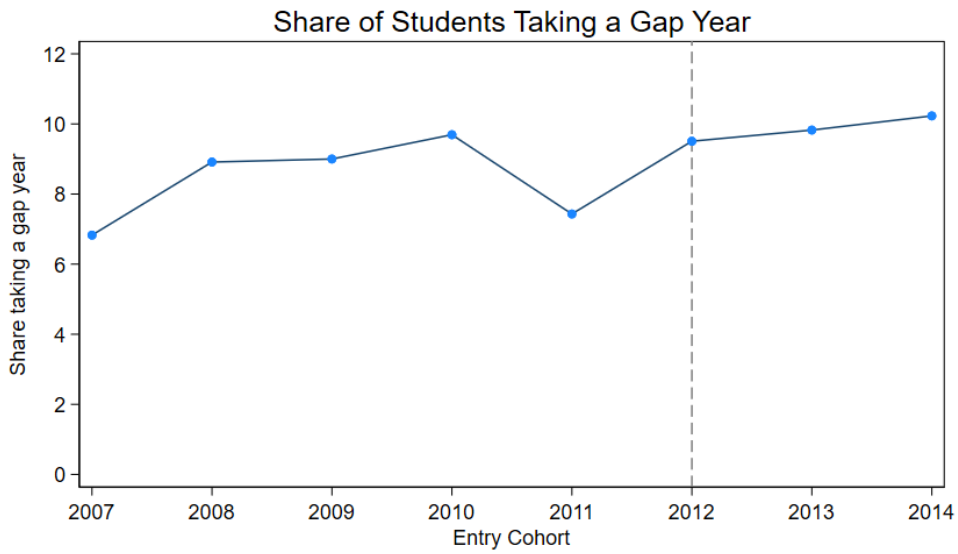


Figure A.1: Share of Students Delaying University Entry by One Year

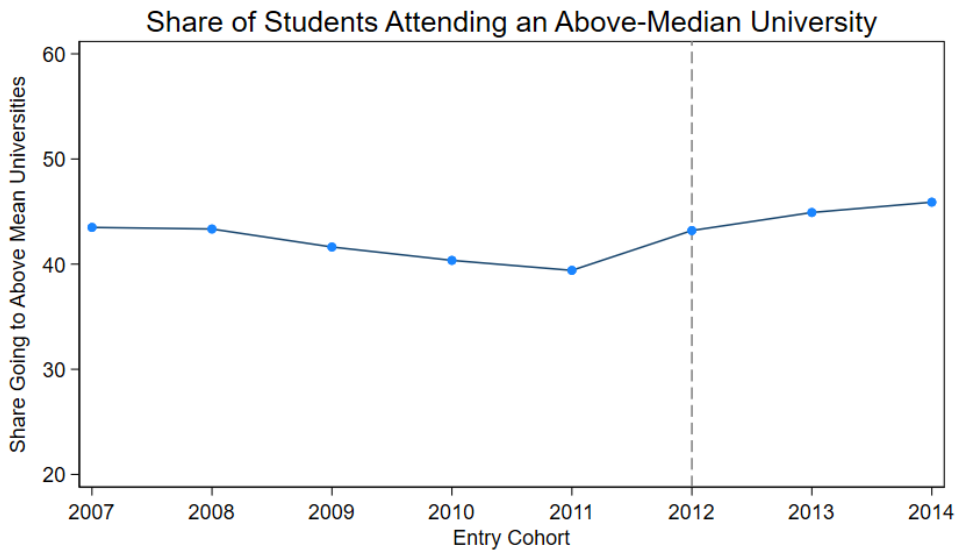


Figure A.2: Share of Students Delaying University Entry by One Year

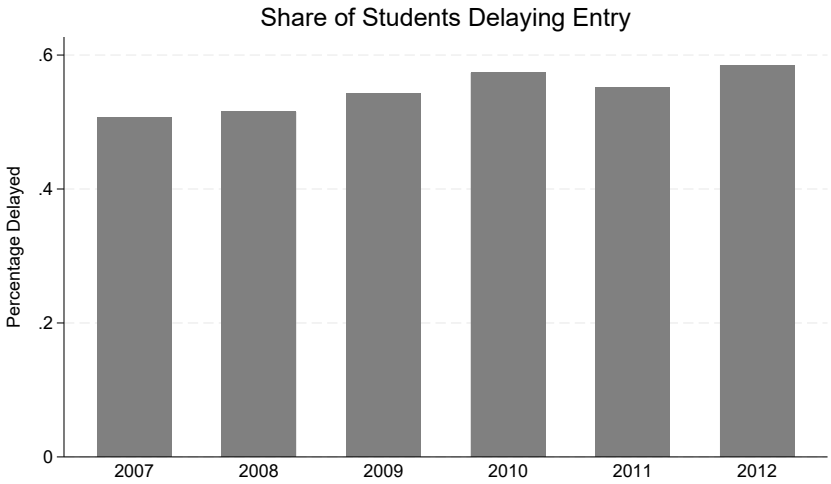
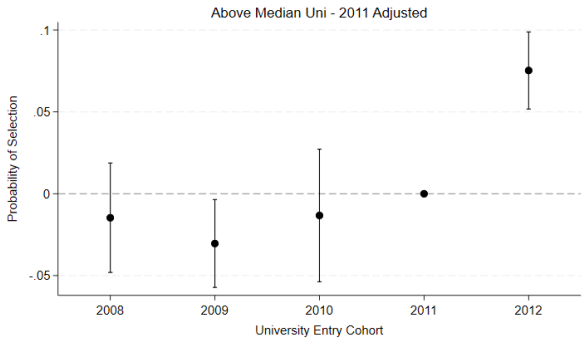
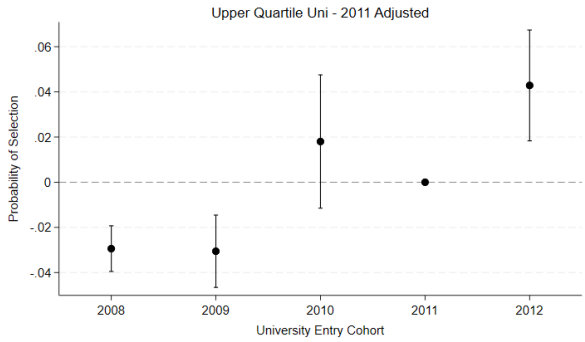


Figure A.3: Figure shows the proportion of students who delay school entry by one year by cohort across the sample.

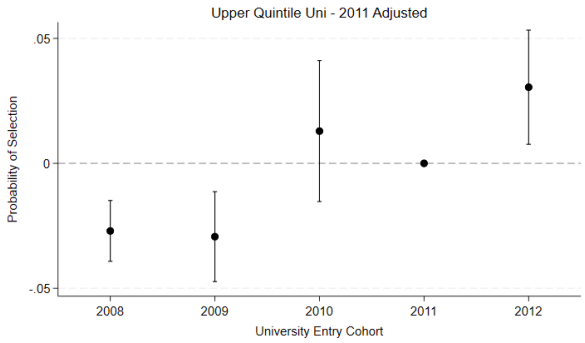
Figure A.4



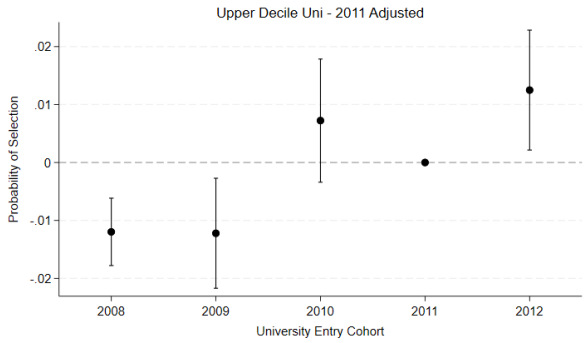
(a) Above-Median institutions



(b) Top quartile



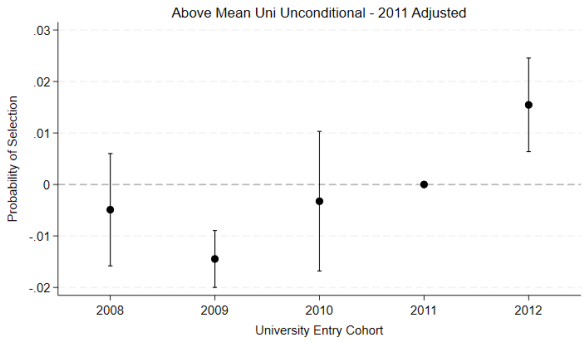
(c) Top quintile



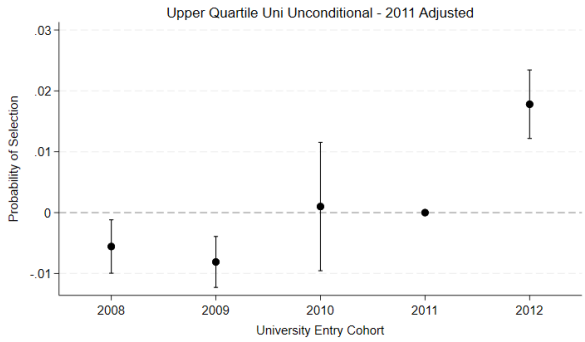
(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on unconditional ever-entry into each selectivity tier, adjusted to predict the participation of the 2011 cohort. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

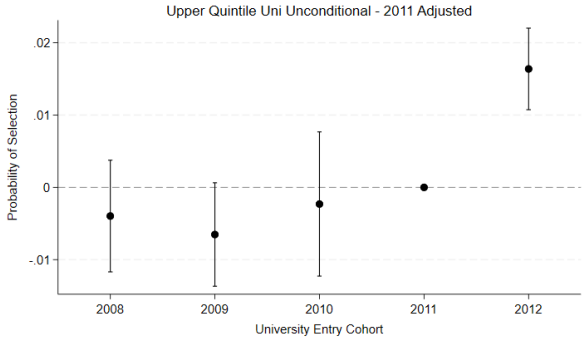
Figure A.5



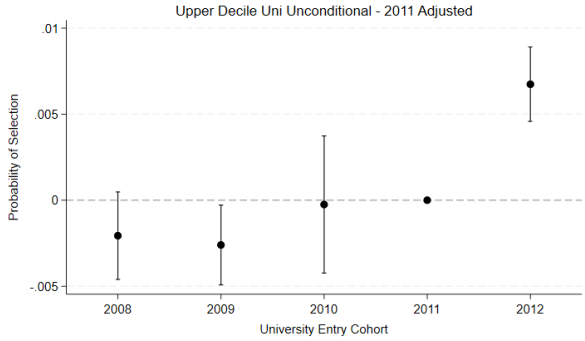
(a) Above-Median institutions



(b) Top quartile



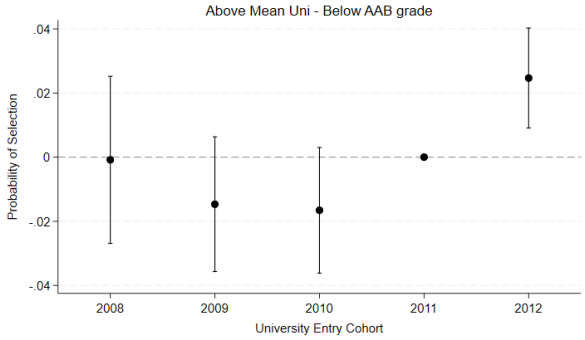
(c) Top quintile



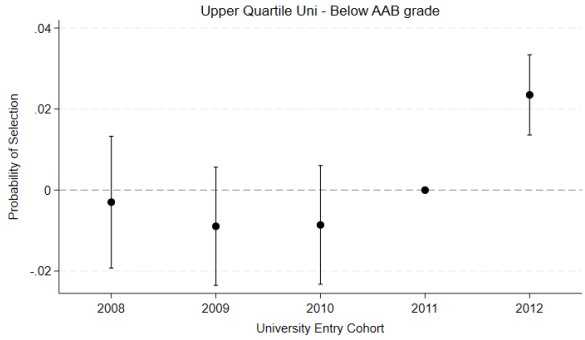
(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on conditional ever-entry into each selectivity tier, adjusted to predict the participation of the 2011 cohort. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

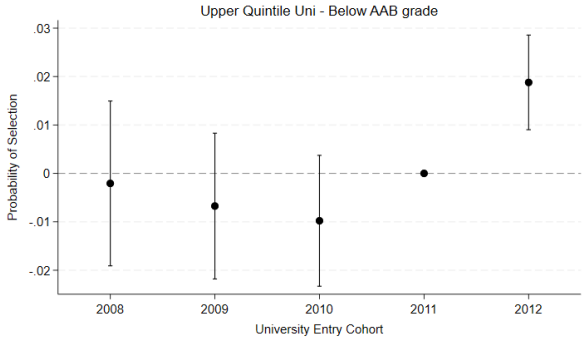
Figure A.6



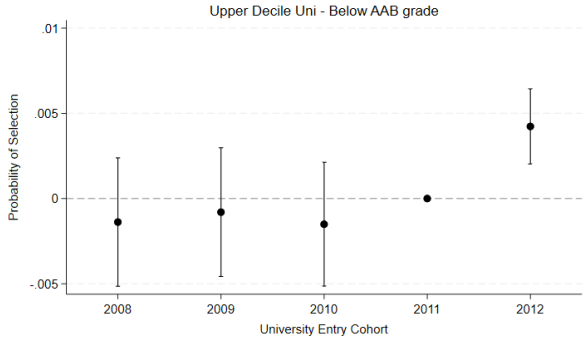
(a) Above-Median institutions



(b) Top quartile



(c) Top quintile



(d) Top decile

Note: Figure plots the estimated difference-in-discontinuity effects on unconditional ever-entry into each selectivity tier for students with A-level attainment below AAB, who were unaffected by the 2012–13 exemption of AAB+ students from student number controls. Data come from the 1998–2020 National Pupil Database and Higher Education Statistical Agency censuses.

Figure A.7

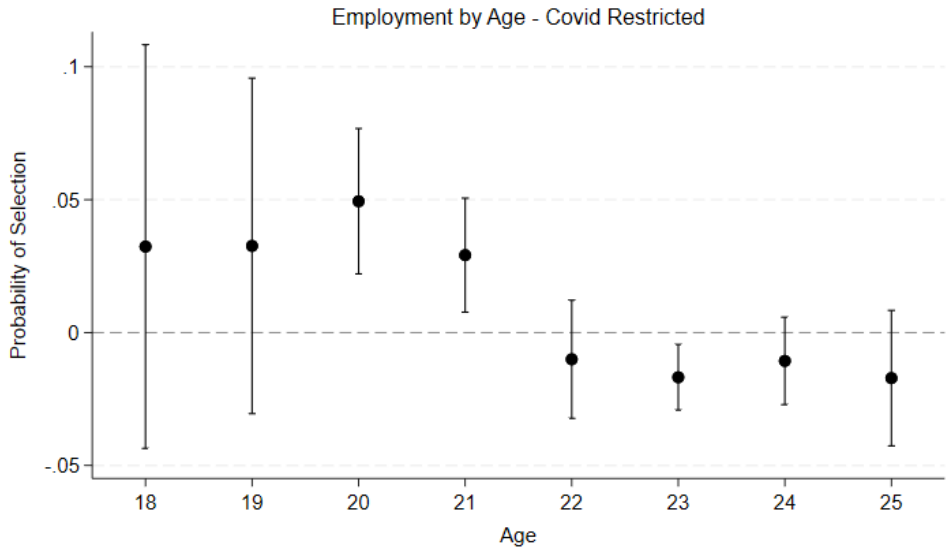


Figure A.8: Note: Figure plots the estimated difference-in-discontinuity effects on employment at each age from 19 to 25 restricted to drop covid affected tax years. Positive estimates at ages 19 and 20 reflect delayed university entrants in work before enrolment; negative estimates at ages 22 and 23 reflect delayed entrants remaining in study while on-time peers entered the labour market; estimates converge to zero by ages 24 and 25. Data come from the 1998–2020 National Pupil Database, the Higher Education Statistical Agency census, and HMRC employment records.

Figure A.9

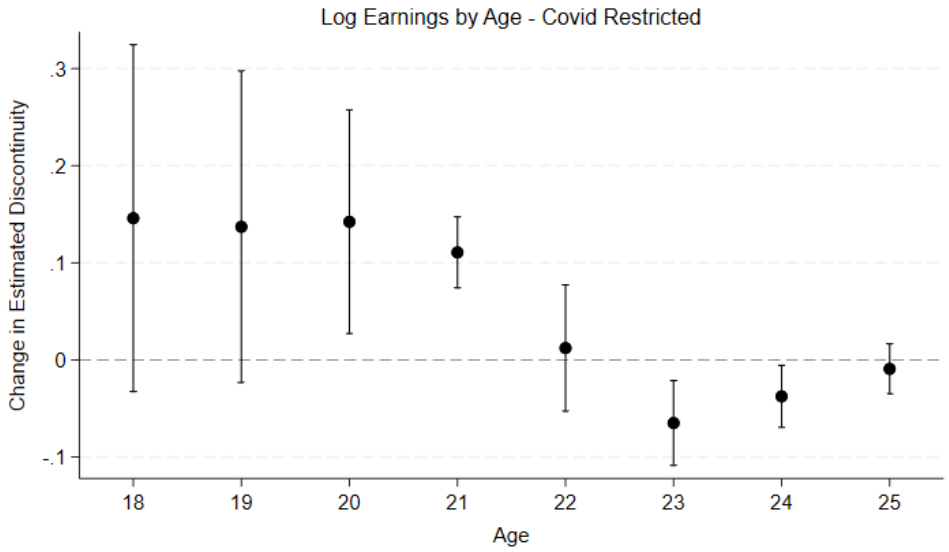


Figure A.10: Note: Figure plots the estimated difference-in-discontinuity effects on earnings at each age from 19 to 25, restricted to drop covid affected tax years. The profile mirrors the employment profile in Figure 10 and matches the entry-timing shifts estimated in Section 6.1. Data come from the 1998–2020 National Pupil Database, the Higher Education Statistical Agency census, and HMRC earnings records.

Table A.1: Difference-in-Discontinuity Estimates by Cohort and Outcome

Note: Each cell reports the difference-in-discontinuity estimate for the cohort column and outcome row. Inference is based on the placebo distribution of pre-reform discontinuities