Predicting College Major Switching with Machine Learning

Yuchen Zhu*

October 26, 2026

Abstract

Major switching can benefit students, but it is costly when it happens late. Advising offices therefore want to know in advance who is likely to switch, so they can help students stay in their major or find a better fit. Using administrative data from the University of Florida, I build a machine learning model that gives each student an approximately unbiased and easily interpretable switching probability. Information available before college predicts switching only weakly, so advising offices cannot reliably identify likely switchers when students first arrive. Prediction improves as students' records accumulate. Performance in a major's core courses is key: it predicts even switches that occur several terms later. At the same advising capacity in each term, the model identifies 18 percent more switchers than the university's current advising system.

*Department of Economics, University of Florida. Email: yuchenzhu@ufl.edu

1 Introduction

Many students choose a field of study before they understand their academic strengths or how well a field suits them. Coursework then provides new information about their interests, abilities, and fit, and they often revise their plans (Malamud, 2011; Bordon and Fu, 2015). At the university I study, about half of students eventually leave the major they start in. This flexibility has value, but switching is costly when it happens late or after a student fails to perform adequately in the major. Completed courses may not count toward the new curriculum (Denice, 2021). Among graduates in my data, students who stay in their starting major until their seventh term or later complete about 19 more credits and take 1.6 more terms to graduate than students who switch in their second term. Late switching also places demands on limited instructional and advising resources and can lower on-time graduation rates.

To avoid late switches, advisors try to reach students early. They review each student's progress and either help the student stay in the current major or advise on which alternative majors to consider. Their time is limited, however, and they must decide which students to prioritize.[1] The relevant question is therefore predictive: given the information available to the university, which students are most likely to switch? Prediction has helped allocate limited resources in settings such as lending, hiring, and public service delivery (Kleinberg et al., 2015; Mullainathan and Spiess, 2017; Kleinberg et al., 2018). Academic advising faces a similar problem.

In this paper, I build a prediction model for major switching using administrative records on 16,605 first-time-in-college students who entered the University of Florida between 2011 and 2015. The records link admissions and demographic information to term-by-term enrollment, major declarations, coursework, grades, credits, and degree completion. Each student is followed for up to 17 terms. At the end of each term, the model predicts switching in

[1]The median caseload for a full-time professional advisor at public doctoral institutions was 285 students (Carlstrom and Miller, eds, 2013). At the University of Florida, the College of Liberal Arts and Sciences reports serving more than 10,000 students with 15 full-time academic advisors.

the following term using only the information available at that point. The records capture students' academic histories and enrollment choices, but not their preferences or intentions to switch, so any predictive power comes from the observed record.

I ask three questions. How well can switching be predicted? What information matters most? And would prediction change which students advisors contact, compared with the university's current advising system?

First, switching can be predicted well. Information available before enrollment yields an AUC of 0.60. Adding time in college, the starting major, and academic performance raises the AUC in a held-out cohort to 0.78 for logit and 0.81 for gradient boosting. The gradient boosting model also produces approximately unbiased switching probabilities that are easy to interpret: among students with a predicted probability near 0.4, about 40 percent switch in the following term.

Second, performance in the starting major's core courses is particularly important. Students who switch later tend to have weaker core-course performance early in college. The importance of other predictors changes as students progress. In the first four terms, predictions rely heavily on academic performance. From term 5 onward, accumulated credits, recent academic setbacks, and characteristics of the starting major receive more weight. Core-course GPA changes little after the first five terms, yet it continues to predict later switching. These patterns show when switching becomes predictable. Without additional data, such as surveys of students' beliefs about majors, an advising office cannot reliably tell who will switch when students first arrive on campus. After a few terms of coursework, it can.

Third, prediction would change which students advisors contact. The current system identifies a set of critical-tracking courses for each major and flags students who fall below thresholds on overall GPA and performance in those courses. The model uses the same signals as continuous measures and adds other information from the academic record. At the same advising capacity in each term, the best-performing model identifies 18 percent

more switchers than the current system. Its predictions also carry information about longer-run switching: predictions made at term 2 reach an AUC of 0.74 for eventual switching. Among students selected by the model but not by the current system, 31 percent switch in the following term and 65 percent eventually leave their starting major.

The targeting comparison assumes that reaching a student who switches in the following term has positive value to the advising office, and that this value is the same across students. Early contact gives advisors an opportunity to discuss students' academic options before additional coursework accumulates. Experimental evidence suggests that advising can improve student outcomes, although effects vary across interventions. Individual coaching has improved persistence and grades (Bettinger and Baker, 2014; Oreopoulos and Petronijevic, 2018), while lighter-touch interventions have shown limited effects (Oreopoulos and Petronijevic, 2019). In a trial of proactive, analytics-informed advising at eleven public research universities, Rossman et al. (2023) find no significant effect on six-year graduation in the pooled sample, but a 7 percentage point increase at Georgia State. The comparison also leaves out potential benefits from contacting students who switch later or remain in their starting major.

This paper contributes to three literatures. First, it extends work on early warning systems in education, which uses administrative and transcript data to identify students at risk of dropping out or falling behind (Agasisti and Bowers, 2017; Baepler and Murdoch, 2010; Picciano, 2012; Sorensen, 2019; Von Hippel and Hofflinger, 2021; Plak et al., 2022), to major switching. Related applications in higher education predict degree completion, eventual major, and program dropout (Beaulac and Rosenthal, 2019; Vaarma and Li, 2024). I focus on switching as a decision made during college and evaluate predictions against an existing advising system, holding advising capacity fixed in each term. Second, the paper relates to research on sequential education and major choice under uncertainty. Education choices unfold as students learn about their abilities, preferences, and expected returns (Altonji, 1993; Arcidiacono, 2004). Work using elicited expectations shows that perceived ability,

expected earnings, and nonpecuniary preferences shape initial major choices (Arcidiacono et al., 2012b; Zafar, 2013; Wiswall and Zafar, 2015). After enrollment, students update their beliefs and intended fields as they learn from college performance (Stinebrickner and Stinebrickner, 2012, 2014; Arcidiacono et al., 2012a). Grades, field-specific achievement, and peer composition are also associated with persistence and switching across fields (Bettinger, 2010; Ost, 2010; Rask, 2010; Astorne-Figari and Speer, 2019). By treating major switching as an out-of-sample forecasting problem, I focus on its timing. Third, the paper adds to the growing use of out-of-sample prediction in applied economics (Kleinberg et al., 2015; Mullainathan and Spiess, 2017; Kleinberg et al., 2018).

Overall, the results give advising offices practical guidance on whom to reach and which signals to watch. Establishing the benefits of reaching these students requires separate evidence on the causal effects of advising, which a staged rollout of the predictions across majors could provide.

2 Data

The analysis uses de-identified administrative records from the University of Florida for 16,605 first-time-in-college students who entered between 2011 and 2015. The sample covers 33 starting majors in Business, Engineering, and Health.[2] The records link admissions and demographic information to term-by-term enrollment, major declarations, coursework, grades, credits, and degree completion. Students are observed for up to 17 academic terms, through graduation or their final observed term. This structure allows me to observe the information available at the end of each term and whether the student changes majors afterward.

The university's current advising system monitors students' progress in their majors during the first five terms. Each major specifies a set of critical-tracking courses required for satisfactory progress, together with GPA thresholds for performance in those courses and

[2]Appendix Table A1 lists the detailed majors in each category.

for overall GPA.

I collect the university's undergraduate catalogs to identify the critical-tracking courses for each major. I refer to these as the major's *core courses*; all other graded coursework is *non-core coursework* relative to the student's starting major.[3] Critical-tracking requirements are concentrated in the first four terms, with some extending into term 5.

The applicable GPA thresholds vary across majors and terms. A student who falls below either threshold is selected for advising and has one term to recover. Students who remain below the requirements in the following term must change majors. I reconstruct selection for advising from the GPA thresholds and students' recorded grades. Actual contact and advising content are unobserved. The reconstructed rule provides the benchmark for evaluating the prediction models in Section 3.

The starting major is the detailed major recorded in a student's first observed term. The forecasting sample is a student-term risk set that includes each term in which the student remains in the starting major and the next term's major is observed. The outcome equals one if the starting major is no longer among the student's declared majors in the following term and zero otherwise. This definition includes moves within the same broad field as well as moves across fields. Students permanently leave the risk set after their first switch or departure from the university. All predictors use information available by the end of the current term.

The predictors combine information available before college with the record that accumulates after enrollment. They include demographic and pre-college characteristics, term in college, characteristics of the starting major, cumulative academic performance, current-term performance and course-taking, and starting-major indicators. Starting-major characteristics include degree credits, STEM status, and median earnings ten years after graduation from the Postsecondary Employment Outcomes (PSEO). Overall GPA and core-course GPA are cumulative measures calculated through the end of the current term.

[3]The classification is major specific, so the same course may be classified as core for one student and non-core for another.

Course easiness is constructed from the grading environment of the courses a student takes. I measure a course's easiness by its mean grade after adjusting for the demographic and pre-college characteristics of the students who took it. Cumulative course easiness is the credit-weighted average across courses completed through the current term. I also measure changes in term GPA and term course easiness relative to the previous term.

For out-of-sample evaluation, the 2011 through 2014 entry cohorts form the development sample used for estimation and tuning, while the 2015 cohort is held out. The cohort split mimics the intended application of using earlier students to predict switching among later entrants.

2.1 Descriptive Patterns

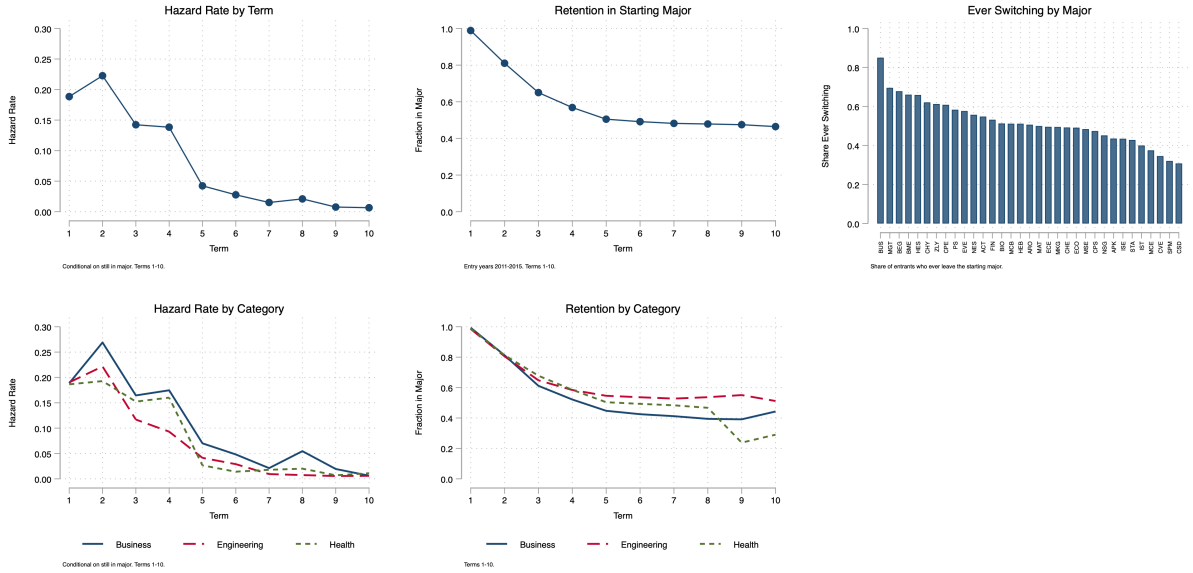
Table 1 reports summary statistics for the forecasting sample. The sample contains 80,895 student-term observations, 12.0 percent of which are followed by a switch in the next term. The switching rate is 12.2 percent in the development cohorts and 11.1 percent in the held-out cohort. Core-course GPA is lower and more dispersed than overall GPA, with means of 3.18 and 3.49, respectively. The development and held-out cohorts have broadly similar college academic records, although the held-out cohort has slightly higher GPAs and fewer D or F grades.

Switching is concentrated early in college (Figure 1). The hazard indexed by term t measures switching in term $t+1$ among students who remain in their starting majors at the end of term t. It is highest in the first four terms and declines sharply thereafter. The share of enrolled students remaining in their starting majors also falls most rapidly early in college. Switching rates vary across starting fields, with Business students leaving at the highest rate and Engineering students at the lowest.

Differences in academic performance appear before many switches occur. Figure 2 compares first-year core-course and non-core GPAs, measured through the end of term 2, by the timing of students' later switches. Students who subsequently switch already have lower

GPAs than those who remain in their starting majors and graduate, and the gap is larger in core courses. The gap is most pronounced among students who switch immediately after the first year and remains visible among later switchers.

Figure 1: Timing and Incidence of Major Switching



Notes: The figure describes switching out of the starting major for the 2011 through 2015 entry cohorts. The switching hazard in term t is the share of students in the starting major at the end of term t who are no longer in it at the end of term $t+1$. The retention series reports, for each term, the share of students enrolled in that term who are still in the starting major. Category-specific panels separate students entering Business, Engineering, and Health majors. The figure is limited to the first 10 terms because the number of students observed declines sharply thereafter and the later-term proportions become unstable.

3 Predicting Major Switching

The empirical analysis essentially consists of estimating a prediction function and then evaluating its performance. For a student who remains in the starting major through term t, the outcome of interest is an indicator for leaving the starting major:

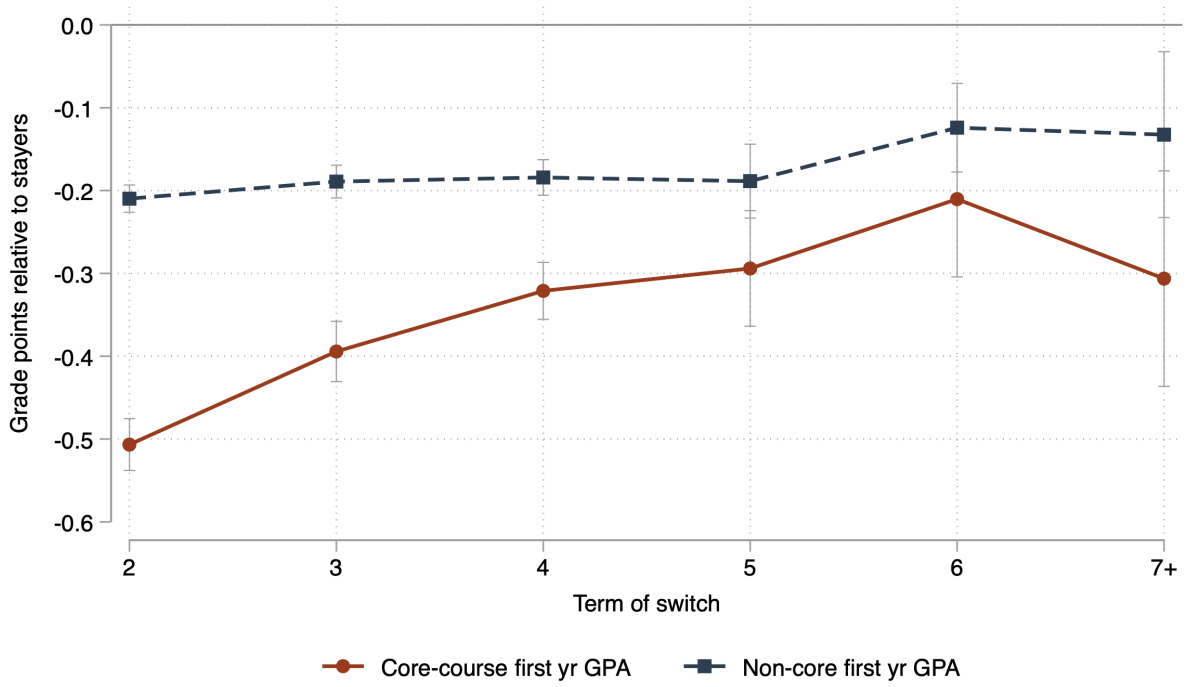
$$Y_{i,t+1} = \mathbf{1}\left\{M_i^0 \notin \mathcal{M}_{i,t+1}\right\},$$

Table 1: Summary Statistics for the Forecasting Sample

	Full Sample 2011–2015	Training Cohorts 2011–2014	Held-Out Cohort 2015
Panel A: Student and Starting-Major Characteristics			
Male	0.505	0.509	0.492
	(0.500)	(0.500)	(0.500)
Underrepresented minority	0.249	0.245	0.260
	(0.432)	(0.430)	(0.439)
First generation	0.081	0.084	0.070
	(0.272)	(0.277)	(0.255)
In-state	0.921	0.931	0.885
	(0.270)	(0.253)	(0.319)
High school GPA	4.24	4.22	4.32
	(0.477)	(0.472)	(0.486)
Math test score	657	657	658
	(81.4)	(83.4)	(73.7)
Verbal test score	637	636	644
	(75.1)	(77.0)	(67.7)
Starting-major credits	122	122	122
	(3.72)	(3.70)	(3.81)
Starting-major STEM intensity	0.897	0.898	0.894
	(0.304)	(0.303)	(0.308)
Starting-major earnings, $1,000s	97.6	97.6	97.8
	(18.1)	(18.1)	(18.1)
Students	16,605	13,003	3,602
Panel B: Student-Term Characteristics			
Next-term switching rate	0.120	0.122	0.111
	(0.325)	(0.328)	(0.314)
Cumulative GPA	3.49	3.48	3.52
	(0.411)	(0.418)	(0.387)
Core-course GPA	3.18	3.18	3.21
	(0.878)	(0.879)	(0.875)
Course easiness	3.33	3.33	3.33
	(0.128)	(0.129)	(0.123)
Change in GPA	-0.028	-0.030	-0.022
	(0.410)	(0.417)	(0.382)
Change in course easiness	0.007	0.006	0.010
	(0.174)	(0.174)	(0.175)
Number of D and F grades	0.076	0.082	0.054
	(0.370)	(0.389)	(0.292)
Term credits	14.7	14.7	14.6
	(3.97)	(3.98)	(3.94)
Student-term observations	80,895	62,906	17,989

Notes: Standard deviations in parentheses. Panel A reports one observation per student, taken at the student's first risk-set term. Changes in GPA and course easiness are set to zero in the first term. High school GPA and test scores are summarized over nonmissing values. Starting-major earnings are median earnings ten years after graduation, in thousands of dollars.

Figure 2: Early Academic Performance by Timing of First Major Switch



Notes: The figure reports mean first-year core-course and non-core GPAs, cumulative through the end of term 2, by the last term the student spends in the starting major. Bars show 95 percent confidence intervals. The vertical distance between the two series captures relative performance inside and outside the core curriculum of the starting major. Group sizes for last terms in the starting major of 2, 3, 4, 5, 6, and 7 or later are 2,379, 1,271, 1,043, 262, 145, and 61.

where M_i^0 denotes the starting major and $\mathcal{M}_{i,t+1}$ is the set of the student's declared majors in term $t+1$. I estimate the one-term-ahead switching probability

$$m(X_{it}) = P(Y_{i,t+1} = 1 \mid X_{it}),$$

where X_{it} contains the input variables and includes only information observed through the end of term t.

The models are estimated on a development sample and evaluated in a separate held-out sample, the 2015 entry cohort. In a standard prediction exercise, evaluation would focus on how well the models predict out of sample. I measure this performance using the area under the receiver operating characteristic curve (AUC). AUC equals the probability that a randomly selected student-term followed by a switch receives a higher predicted risk than one not followed by a switch (Hanley and McNeil, 1982).[4]

The advising application goes beyond overall predictive performance. The model generates a continuous predicted probability, while the university's current advising system uses a binary flag based on academic thresholds to identify students for advising contact. The flag provides a simple operational rule when the advising office can contact only a subset of students each term.[5] I therefore take the university's current advising system as the institutional baseline. During the first five terms, the system flags about 17 percent of student-term observations in the evaluation cohort. Let K_t denote the number of students flagged in term t. At this capacity, selecting the K_t students with the highest predicted switching probabilities maximizes the expected number of next-term switchers reached under the equal-value assumption. I compare this ranking with the current system to assess whether it identifies more switchers with the same number of contacts.

[4]Other measures capture different aspects of predictive performance. Classification accuracy depends on a particular classification threshold and can be heavily influenced by the more common outcome when switching is infrequent. AUC is less sensitive to this class imbalance because it compares the relative ranking of switchers and non-switchers. Calibration asks whether predicted probabilities correspond to observed switching rates and is reported in Figure A1. The allocation exercise depends mainly on the ordering of students by predicted risk, and AUC evaluates this ordering across classification thresholds (Fawcett, 2006).

[5]Contact refers to advising outreach, such as an advising session or an email reminder.

The costs associated with later switching give a reason to identify likely switchers early. Among graduates, later switching is associated with more credits and more time to degree (Appendix Table A4). Compared with students who stay in their starting major until term 2, those who stay until term 6 complete about 8 additional credits. Those who stay until term 7 or later complete about 19 additional credits and take 1.6 more terms to graduate. Valuing the 19 additional credits at UF's resident tuition and fee rate gives a tuition equivalent of approximately $4,100.[6] Earlier identification may give students and advisors more time to review major options before additional coursework accumulates.

Following the logic of Kleinberg et al. (2018), I evaluate the model against the current advising system in two complementary ways. First, I hold the contact count at K_t in each term and compare how many next-term switchers each approach identifies. Second, I hold coverage at the current system's level and ask how many contacts the model needs to identify the same number of switchers. These comparisons measure targeting efficiency while holding observed switching outcomes fixed. Their implications for student outcomes depend on how advising affects students' decisions.

3.1 Predictive Content of the Academic Record

I first examine how much predictive information each part of the academic record provides. Using the same logit functional form, I estimate a sequence of nested specifications that add one group of predictors at a time. The sequence begins with pre-college performance and then adds demographics, term fixed effects, starting-major fixed effects, college GPA measures, and current-term performance. Changes in held-out AUC show how much each new block improves prediction beyond the information already included.

Predictive performance improves substantially only after students enter college (Table 2). Pre-college academic performance provides only modest predictive information, with an AUC of 0.596 compared with 0.5 under random guessing. Demographics add essentially no pre-

[6]The calculation uses UF's 2026/27 undergraduate resident tuition and fee rate of $214.54 per credit hour. It excludes living expenses and the opportunity cost of additional time in college.

dictive value, leaving the AUC nearly unchanged at 0.597. Once college records become available, performance improves sharply. Accounting for how far students are into their college careers raises the AUC to 0.729, and starting-major differences raise it further to 0.745. Adding college GPAs brings the AUC to 0.781, helping distinguish switchers from non-switchers within the same term and starting major. Current-term performance adds little beyond this, raising the AUC to 0.782. Because the specifications are cumulative, a small increase does not imply that newly added variables are uninformative on their own, since some of their predictive content may already be captured by variables entered earlier. The sequence shows that most of the predictive gain emerges after enrollment, particularly once students' stage in college and cumulative academic performance become observable.

Table 2: When Does Information Become Predictive?

Information available through term t	Next-term switching AUC
(1) Pre-college performance	0.5960
(2) + Demographics	0.5974
(3) + Term fixed effects	0.7292
(4) + Major fixed effects	0.7449
(5) + College GPAs	0.7808
(6) + Current-term performance	0.7818

Notes: Entries report AUC in the 2015 evaluation cohort for predicting first departure from the starting major in the following term. Specifications are cumulative, so each row adds the listed information to all variables included above it. Pre-college performance includes high-school GPA and standardized test scores. Demographics add gender, race and ethnicity, first-generation status, and in-state status. Term fixed effects add indicators for term in college. Major fixed effects add indicators for the student's starting major. Major-level attributes are omitted from the build-up because they are absorbed once major fixed effects are included. College GPA adds cumulative overall GPA, cumulative core-course GPA, and cumulative course easiness. Current-term performance adds Δ GPA, Δ course easiness, the number of D/F grades in the current term, and current-term credits. Row-to-row changes in AUC are conditional on the information already included and therefore depend on the ordering of the blocks; they do not measure each block's standalone predictive value. The large increase from adding term fixed effects primarily reflects differences in switching rates across terms.

The same observed information can also be used to study conditional associations with switching, but the object of interest is different. The prediction analysis focuses on the fitted

value, $\hat{y}_{it}$, which combines the observed characteristics into a predicted probability of switching. Its performance depends on how well that score ranks students out of sample rather than on the precision of any individual coefficient (Shmueli, 2010). The regression analysis instead focuses on estimated coefficients, which describe how individual characteristics are associated with next-term switching conditional on the rest of the observed record. A characteristic may therefore have a precisely estimated conditional association with switching while adding little to the ordering of $\hat{y}_{it}$ once related information is already contained in the fitted score.

Turning to the logit estimates, academic performance is strongly associated with next-term switching. The average marginal effect of overall GPA is -8.6 percentage points, against a base switching rate of 12.2 percent (Table 3). The corresponding marginal effect for the number of D or F courses in the current term is 2.4 percentage points. High school GPA also retains a negative conditional association after college performance is observed, with an average marginal effect of -2.3 percentage points.

The persistence of the high-school GPA association is consistent with evidence that pre-college grades remain informative about college outcomes beyond the first year (Geiser and Santelices, 2007; Allensworth and Clark, 2020). The progressive AUC results show that pre-college performance alone provides modest predictive information. More generally, association and prediction are distinct: a variable may remain conditionally related to switching without changing the ordering of students very much once the rest of the academic record is known.

3.2 Improving Risk Ranking with Flexible Prediction

The logit model provides a useful benchmark, but its functional form may miss nonlinearities and interactions that are useful for prediction. I therefore compare it with several machine-learning methods that use the same observed information but allow it to enter the prediction more flexibly. This comparison shows whether greater model flexibility improves

Table 3: Selected Academic Signals in the Logit

Predictor	Average marginal effect	Mean	SD	One-SD change (pp)
High school GPA	−0.0229***	4.22	0.47	−1.1
	(0.0053)			
Overall GPA	−0.0860***	3.48	0.42	−3.6
	(0.0040)			
Cumulative core-course GPA	−0.0139***	3.18	0.88	−1.2
	(0.0015)			
Cumulative course easiness	0.0821***	3.33	0.13	1.1
	(0.0119)			
D or F courses in the current term	0.0239***	0.08	0.39	0.9
	(0.0034)			
Credits in the current term	−0.0016***	14.7	3.98	−0.6
	(0.0004)			

Average marginal effects are changes in next-term switching probability, with standard errors in parentheses; the base rate in the development cohorts is 12.2 percent. The last column multiplies the average marginal effect by the standard deviation and reports the result in percentage points. Means and standard deviations are for the development cohorts over student-term observations, except high school GPA, which is over students. The specification includes demographics, pre-college measures, term indicators, college GPA measures, detailed course performance, and starting-major fixed effects. The full specification is reported in Appendix Table A3. $^{***}p < 0.01$.

out-of-sample predictive performance. LASSO regularizes the linear specification (Tibshirani, 1996), while random forests (Breiman, 2001) and gradient-boosted trees (Friedman, 2001) allow for nonlinearities and interactions without specifying them in advance. Gradient boosting is implemented with XGBoost (Chen and Guestrin, 2016). Hyperparameters are selected by five-fold cross-validation on AUC within the development sample.[7] All models use the same predictor set and the same held-out evaluation cohort.

XGBoost performs best in the held-out cohort, with an AUC of 0.811 compared with 0.782 for the logit benchmark (Table 4). The 0.029 gain is meaningful in this setting: it exceeds the improvement from adding starting-major fixed effects to the logit in Table 2 (0.0157), though it remains smaller than the gain from adding college GPAs (0.0359). Much of the predictive signal is therefore already captured by the additive logit specification, while nonlinearities and interactions further improve the ranking. XGBoost is therefore used for

[7]Folds are grouped by student, so no student contributes observations to both a training and a validation fold.

the remaining analyses.

XGBoost is also well calibrated in the held-out cohort, with predicted probabilities closely tracking observed switching rates across risk deciles (Appendix Figure A1). The probabilities therefore have an interpretable scale: among students with a predicted switching probability near 0.4, about 40 percent switch in the following term.

AUC orders the models, but it does not show whether a better ranking identifies more switchers at the current system's contact capacity. I therefore compare XGBoost with the university's current advising system at its term-specific contact capacity, K_t. In each of the first five terms, XGBoost selects the same number of students from the same risk set of students who have not yet switched. Across the five terms, this amounts to 2,214 contacts in the evaluation cohort.

Figure 3 extends the comparison beyond the current system's capacity by scaling contacts to λK_t in each term, with $\lambda = 1$ corresponding to its capacity. At $\lambda = 1$, the vertical difference gives the additional switchers identified by XGBoost. At the current system's coverage level, the horizontal difference gives the reduction in contacts needed to identify the same number of switchers.

At the current system's capacity, XGBoost identifies 42.9 percent of next-term switchers, compared with 36.3 percent under the current system, an 18 percent relative increase. The improvement appears in each of the first five terms, as reported in Table A5. XGBoost can also match the current system's coverage with about 24 percent fewer contacts. A logit ranking using the same predictors identifies 38.7 percent of next-term switchers. The difference between XGBoost and logit accounts for roughly two thirds of XGBoost's gain over the current system. Table 5 reports the full comparison.

The model also identifies students the current advising system misses. Figure A2 shows that much of the difference in student selection occurs around the GPA thresholds. Students selected only by XGBoost often have core-course GPAs just above the cutoff, while those selected only by the current system tend to fall just below it. The model thus reallocates

advising attention among students with similar academic performance by using additional information about their switching risk. Of the 811 next-term switchers identified by XGBoost, 232 (29 percent) were never flagged by the current system before switching.

The next-term predictions also identify longer-run switching risk (Table 6). In the held-out cohort, the predicted probabilities at term 2 have an AUC of 0.731 for next-term switching and 0.739 for eventual switching. A model trained directly to predict eventual switching achieves an AUC of 0.746. The next-term model therefore provides a ranking of longer-run switching risk close to that obtained by predicting eventual switching directly.

Among students identified only by XGBoost in a given term, 31 percent switch in the following term and 65 percent eventually switch out of their starting major (Figure 4). The corresponding rates among students identified only by the current system are 17 and 50 percent. About half of the students in the XGBoost-only group who remain in their starting major in the following term eventually switch.[8]

One thing to note is that these comparisons evaluate targeting, not the causal effect of advising. Realized switching may itself depend on advising contact. Let $Y_{i,t+1}(1)$ denote whether student i switches in the following term after receiving contact under the current advising system in term t, and let $Y_{i,t+1}(0)$ denote the corresponding outcome without that contact, holding prior contact histories fixed. Students identified only by XGBoost do not receive contact under the current system and reveal $Y_{i,t+1}(0)$, whereas those identified only by the current system receive contact and reveal $Y_{i,t+1}(1)$. Advising may help some students remain in their starting major, while encouraging others to switch. Because these groups contain different students, their observed switching rates do not identify the causal effect of advising.

Nevertheless, a sensitivity calculation helps assess how much current-term advising contact would have to reduce switching to eliminate the observed targeting gap (Guerdan et al.,

[8]The comparison pools student-term observations from terms 1 through 5 of the held-out 2015 cohort. Groups are defined by whether each method identifies a student in a given term and held fixed when examining later outcomes. Eventual switching is measured through the end of follow-up.

2024). Each method identifies 894 student-term observations that the other method does not identify. The next-term switching rate is 16.7 percent among students identified only by the current system and 30.6 percent among those identified only by XGBoost (Table 7). Holding prior contact histories and the XGBoost-only switching rate fixed, eliminating the gap would require the current-system-only group's switching rate without current-term contact to reach 30.6 percent, approximately 14 percentage points above its observed rate. Under this scenario, contact would have prevented 125 of 274 switches, a reduction of about 46 percent.

Table 4: Predictive Performance across Models

	Logit	LASSO	Random forest	XGBoost
Development-sample AUC	0.793	0.793	0.836	0.865
Evaluation-cohort AUC	0.782	0.782	0.796	0.811

The outcome is whether a student leaves the starting major in the following term, with the risk set restricted to students who have not yet switched. All models use the same predictor set. For the penalised and tree-based models, hyperparameters are selected by five-fold cross-validation within the development cohorts, with folds grouped by student so that repeated observations on the same student do not appear on both sides of a fold. Development-sample AUC is calculated on the same observations used to estimate each model. Evaluation uses the held-out 2015 entry cohort.

Table 5: Targeting Performance Relative to the Advising Rule

	Contact rule	XGBoost	Logit	Random
Contacts	2,214	2,214	2,214	2,214
Switchers among those contacted	686	811	731	354
of which: never flagged by the rule	0	232	248	202
of which: without two prior below-threshold terms	367	555	538	307
Share of contacts who switch	0.310	0.366	0.330	0.160
Share of all switchers contacted	0.363	0.429	0.387	0.187

Notes: 2015 holdout cohort, terms 1–5. The unit is the student-term: a student can be contacted in more than one term. Each policy contacts K_t students per term, where K_t equals the number flagged by the rule in that term's risk set. Random is the analytic expectation at the same contact rate. Never flagged refers to students the rule does not flag in any term. Not preceded by two below-threshold terms refers to switches that do not follow two consecutive terms below the rule's thresholds, a proxy for switches the threshold itself does not force.

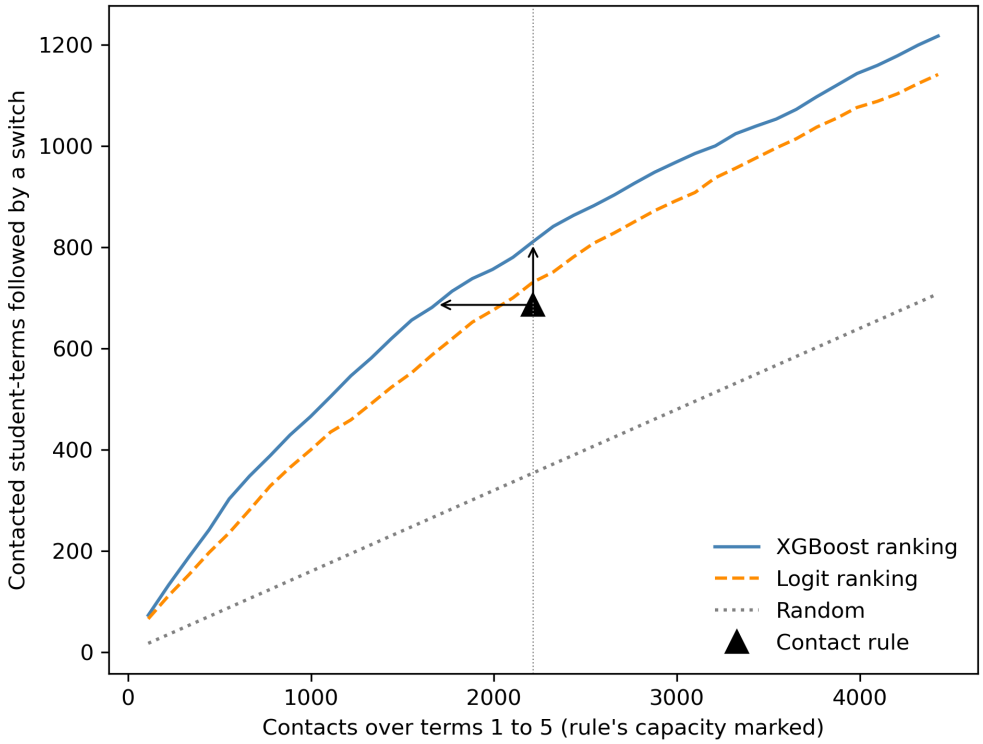


Figure 3: Switchers Reached against Contacts: the Ranking and the Rule

Notes: 2015 evaluation cohort, terms 1 through 5, risk set of students who have not yet left the starting major. The vertical axis is the number of contacted student-terms followed by a switch in the next term. The horizontal axis is total contacts over the five terms. Each point on a curve corresponds to a scalar λ: every term contacts λK_t student-terms, where K_t is the number the rule flags in term t, and $\lambda = 1$ is the rule's own capacity of 2,214 contacts, marked by the vertical line. The XGBoost and logit curves contact the λK_t highest-scoring student-terms in each term; the dotted line contacts λK_t student-terms at random. The rule is the single point at $\lambda = 1$. The vertical arrow is the additional switchers reached at the rule's capacity (811 against 686); the horizontal arrow is the reduction in contacts needed to reach the rule's 686 switchers.

Table 6: Predictive Performance over Longer Horizons

Term	Next term	Within 2 terms	Within 3 terms	Eventual switching
1	0.673	0.701	0.712	0.701
2	0.731	0.744	0.743	0.739
3	0.749	0.752	0.763	0.748
4	0.750	0.769	0.765	0.745

Notes: Entries report AUCs for XGBoost in the held-out 2015 cohort. Within each term, the same next-term predicted probabilities are evaluated over different outcome windows. AUC is calculated separately within each term. The two- and three-term windows include switching within the next two or three terms, respectively. Eventual switching is measured through the end of follow-up. At term 2, a model trained directly to predict eventual switching achieves an AUC of 0.746.

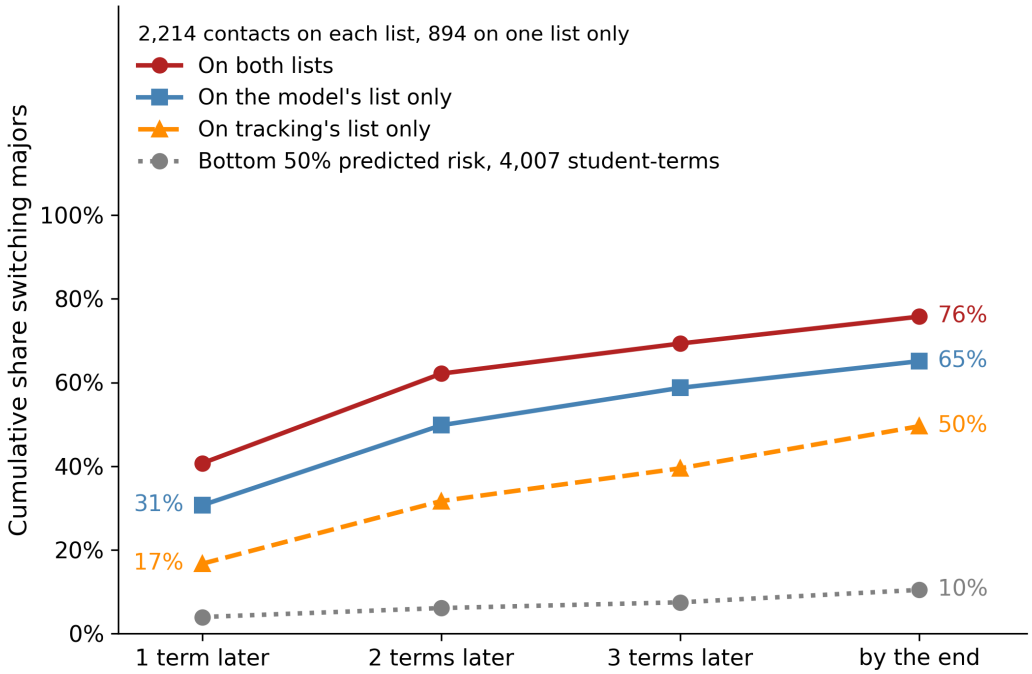


Figure 4: Switching over Longer Horizons

Notes: The figure pools student-term observations from terms 1 through 5 of the held-out 2015 cohort. Each method identifies 2,214 student-term observations at the same term-specific capacity, with 1,320 identified by both methods and 894 identified by each method alone. Groups are defined by whether each method identifies a student in a given term and held fixed when examining later outcomes. The figure reports cumulative switching out of the starting major over the following one, two, and three terms and through the end of follow-up.

Table 7: Sensitivity to Advising Contact

Group	Student-terms	Next-term switches	Switching rate
Both methods	1,320	537	40.7%
Current system only	894	149	16.7%
XGBoost only	894	274	30.6%

Notes: The table uses the held-out 2015 cohort, terms 1 through 5. Both methods use the current system's term-specific contact capacity, totaling 2,214 student-term observations per method. Groups are defined by whether each method identifies a student in a given term. Switching rates refer to departure from the starting major in the following term.

3.3 Predictive Information

Unlike a logit model, XGBoost does not produce coefficients that summarize how each predictor is associated with switching risk. To identify which features contribute most to its predictions, I use SHAP values (Lundberg and Lee, 2017), reported in Figure 5.[9] SHAP decomposes each fitted prediction into contributions from the individual predictors, measured relative to the model's average prediction. A positive SHAP value indicates a contribution toward higher predicted switching risk, while a negative value indicates a contribution toward lower risk. Its magnitude measures the strength of that contribution.[10] For student-term observation it,

$$f(X_{it}) = \phi_0 + \sum_{p=1}^{P} \phi_{pit},$$

where ϕ_0 is the baseline model output and ϕ_{pit} is the contribution of predictor p to the fitted value for observation it.

The SHAP results show that time in college is the largest contributor to the pooled model's predictions, reflecting differences in switching risk across terms. Core-course and overall GPA are the strongest student-level signals and help distinguish switchers from non-switchers within the same term. Lower overall GPA and lower performance in the starting major's core curriculum are associated with higher predicted switching risk, while stronger performance is associated with lower risk.

The model also picks up information from students' course-taking patterns. A lighter course load in the current term is associated with higher predicted switching risk. Students who have taken relatively easier courses also tend to have higher predicted risk, although the relationship is nonlinear. Together, these patterns suggest that course-taking contains predictive information beyond realized academic performance. Characteristics of the starting

[9] Appendix Figure A3 reports the corresponding feature ranking based on mean absolute SHAP values.

[10] SHAP values measure each feature's contribution to the model output relative to the baseline prediction. Their magnitude should therefore be interpreted as the strength and direction of a feature's contribution, not as a percentage-point change in switching probability.

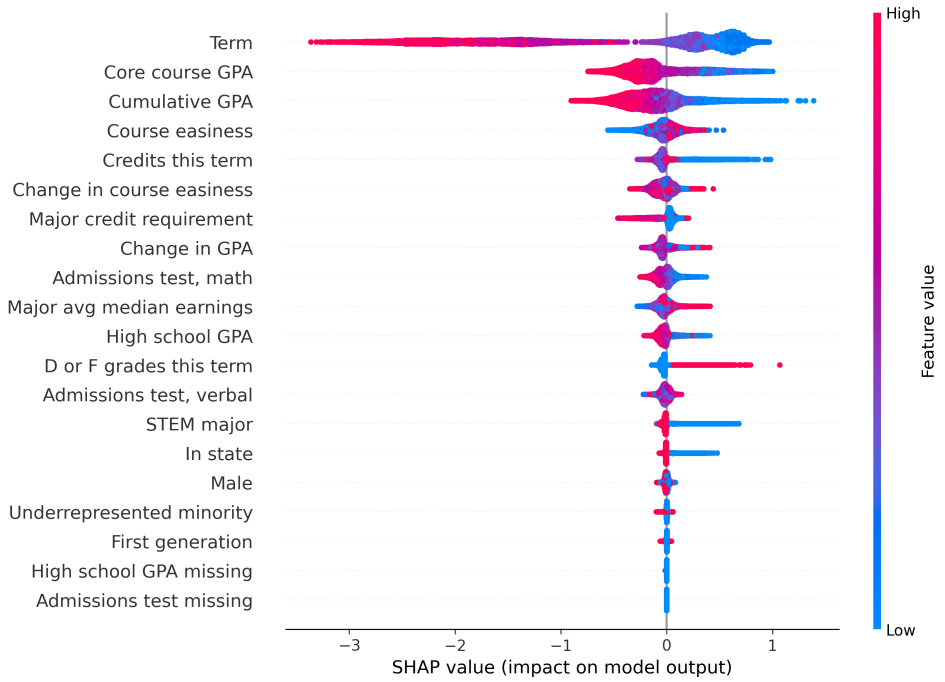


Figure 5: SHAP Summary for the Timely XGBoost Model

Notes: Each point is a student-term in the 2015 evaluation cohort. Starting-major indicators remain in the fitted model and are omitted from the displayed ranking for readability.

major and academic term also contribute to the prediction.

3.4 Predictive Information Across Groups

The information underlying the pooled model's predictions may differ across groups. I examine three dimensions that are especially relevant for advising: stage in college, standing relative to the advising GPA thresholds at term 1, and starting field.

The stage split separates terms 1 through 4 from term 5 and later. Switching rates are 15 to 25 percent of the risk set per term in the first four terms, then fall to 1 to 4 percent afterward (see Figure 1). This split also separates the earlier period, when switching is relatively common, from the later period, when switching is rarer and, as shown earlier, more costly. The standing split separates students who were below at least one advising GPA threshold at term 1 from those who met both thresholds (Appendix Figure A4). The split isolates students who performed relatively well at the outset, for whom early academic

difficulty was less apparent. The field split compares Business, Engineering, and Health (Appendix Figure A5).

For each comparison, I fit a separate XGBoost model for each group using that group's observations in the development cohorts and the hyperparameters of the pooled model. I then compute SHAP values on the corresponding observations in the 2015 evaluation cohort. To compare the information used across groups, I express each predictor's mean absolute SHAP value as a share of the total across predictors within that group. The shares sum to one, so they describe the composition of the ranking rather than the overall magnitude of the SHAP contributions.[11] Figure 6 reports the stage comparison.

Across all three comparisons, core-course GPA and overall GPA are the two largest contributors. What changes is how much of the ranking they account for and which predictors make up the remainder. These differences are largest across stages of college, smaller across term-1 standing, and smallest across starting fields.

Early in college, the ranking is dominated by academic performance. Core-course GPA and overall GPA together account for half of total mean absolute SHAP. Their combined share falls substantially in the late risk set, driven mainly by the decline in core-course GPA. Because core coursework is largely complete by then, this measure mainly carries forward information about how the student performed earlier in college.

The remaining late-stage contribution shifts toward characteristics of the major and the student's current term. Major requirements become more important, especially the number of credits required for the major, whose share rises from 0.04 to 0.10. Current-term performance and course load also carry more weight, while pre-college measures become somewhat less important. Although core-course GPA no longer changes after term 5, it remains predictive of later switching. The late ranking therefore combines earlier core-course performance with more recent academic information and starting-major characteristics. For students beyond term 5, when the current system's GPA rule no longer applies, these signals

[11] Mean absolute SHAP summarizes the magnitude of each feature's contribution but does not show how predicted risk varies over the feature's range.

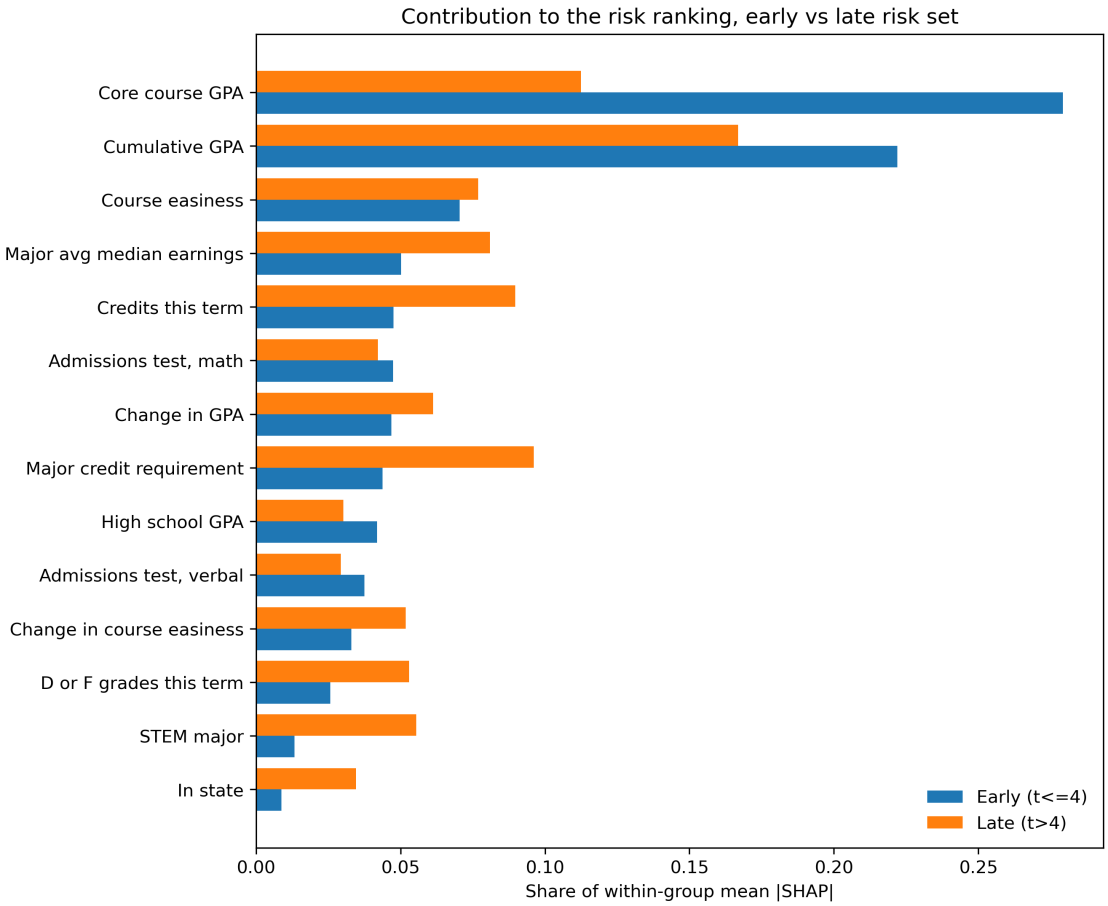


Figure 6: Predictive Contributions across the College Career

Notes: Each bar reports a predictor's mean absolute SHAP value as a share of the total across predictors within the group. Separate models are fitted for early and late observations in the development cohorts using the hyperparameters of the pooled model. SHAP values are computed on the corresponding observations in the 2015 evaluation cohort. Early observations cover terms 1 through 4 and late observations cover term 5 and later. Starting-major indicators remain in the fitted models and are omitted from the display. Predictors shown have a share of at least 0.03 in at least one group. Bars are sorted by the early group.

may help identify those who remain at risk of switching.

The standing split shows a similar pattern. For students who began below an advising GPA threshold, the prediction relies more heavily on academic performance, especially overall GPA. Among students who initially met the thresholds, the model draws relatively more on recent changes in performance, pre-college measures, and characteristics of the starting major.

Differences across starting fields are smaller. In Health, overall GPA has the largest share at 0.30. In Engineering, core-course GPA is largest at 0.23, with relatively more contribution from the admissions math test and from a D or F in the current term. Business distributes predictive contribution more evenly across the observed predictors. Despite these differences, the two GPA measures remain the leading contributors in each field.

Differences in the underlying risk sets may also contribute to these patterns. The late group and the students who initially met the advising GPA thresholds have higher and less dispersed GPAs than their counterparts, leaving less variation in GPA for the model to use. The late risk set also contains fewer switchers, so its SHAP profile should be interpreted more cautiously. Across all three splits, however, the same two GPA measures remain the leading contributors. The models therefore draw on a common set of academic signals, while the relative contribution of other information changes most over the college career.

4 Conclusion

This paper asks how well university records predict major switching and whether those predictions can improve the allocation of limited advising attention. Pre-college information provides only modest predictive value. Once college records become available, XGBoost achieves a held-out AUC of 0.81 and produces approximately unbiased switching probabilities. At the same contact capacity in each term, it identifies 18 percent more next-term switchers than the current advising system. Predictive value also extends to longer horizons:

the next-term model's predictions at term 2 achieve an AUC of 0.74 for eventual switching. Among students identified only by XGBoost who remain in their starting major in the following term, about half eventually switch out.

Time in college captures differences in switching risk across terms, while core-course and overall GPA help distinguish students within the same term. Core-course GPA remains predictive after it stops changing at the end of term 5, and later predictions draw on a broader set of academic and major characteristics.

The value of contacting a student depends on whether switching would improve that student's outcomes and on which advising response would help most. Establishing these benefits requires separate evidence on the causal effects of advising. A natural next step is to roll out the predictions major by major, so that advisors in some majors use them to guide outreach before others do. Majors that have not yet adopted the predictions would then serve as a comparison group. Changes in major choices, credit accumulation, and progress toward graduation in adopting majors, relative to that group, would identify the effect of prediction-guided outreach. Because the model selects students by predicted switching risk, advisors using it would also contact students who are above the current performance thresholds. Outreach would then no longer signal poor performance, which may reduce the discouragement that a warning can cause. In related work on the university's GPA requirements, Ainsworth et al. (2026) find that discouragement is a key reason why falling below a threshold lowers graduation.

The scope of these findings depends on the institutional setting. The estimates come from the University of Florida, where critical-tracking coursework is concentrated early in college and the analysis focuses on students entering Business, Engineering, and Health. These features may shape both switching patterns and the predictive content of academic records.

References

Agasisti, Tommaso and Alex J Bowers, "Data analytics and decision making in education: towards the educational data scientist as a key actor in schools and higher education institutions," in "Handbook of contemporary education economics," Edward Elgar Publishing, 2017, pp. 184–210.

Ainsworth, Robert, Thomas Knight, Macy McClaren, and Yuchen Zhu, "The Academic Effects of Performance Requirements for Staying in College Majors," 2026. Working paper.

Allensworth, Elaine M and Kallie Clark, "High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools," *Educational Researcher*, 2020, *49* (3), 198–211.

Altonji, Joseph G, "The demand for and return to education when education outcomes are uncertain," *Journal of Labor Economics*, 1993, *11* (1, Part 1), 48–83.

Arcidiacono, Peter, "Ability sorting and the returns to college major," *Journal of econometrics*, 2004, *121* (1-2), 343–375.

_ , Esteban M Aucejo, and Ken Spenner, "What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice," *IZA Journal of Labor Economics*, 2012, *1* (1), 5.

_ , V Joseph Hotz, and Songman Kang, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

Astorne-Figari, Carmen and Jamin D Speer, "Are changes of major major changes? The roles of grades, gender, and preferences in college major switching," *Economics of Education Review*, 2019, *70*, 75–93.

Baepler, Paul and Cynthia James Murdoch, "Academic analytics and data mining in higher education," *International Journal for the Scholarship of Teaching and Learning*, 2010, *4* (2), 17.

Beaulac, Cédric and Jeffrey S Rosenthal, "Predicting university students' academic success and major using random forests," *Research in Higher Education*, 2019, *60* (7), 1048–1064.

Bettinger, Eric, "To be or not to be: Major choices in budding scientists," in "American universities in a global market," University of Chicago Press, 2010, pp. 69–98.

Bettinger, Eric P. and Rachel B. Baker, "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising," *Educational Evaluation and Policy Analysis*, 2014, *36* (1), 3–19.

Bordon, Paola and Chao Fu, "College-major choice to college-then-major choice," *The Review of economic studies*, 2015, *82* (4), 1247–1288.

Breiman, Leo, "Random Forests," *Machine Learning*, 2001, *45* (1), 5–32.

Carlstrom, Aaron H. and Marsha A. Miller, eds, *2011 NACADA National Survey of Academic Advising* Monograph No. 25, Manhattan, KS: National Academic Advising Association, 2013.

Chen, Tianqi and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," in "Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" 2016, pp. 785–794.

Denice, Patrick A, "Choosing and changing course: Postsecondary students and the process of selecting a major field of study," *Sociological perspectives*, 2021, *64* (1), 82–108.

Fawcett, Tom, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, 2006, *27* (8), 861–874.

Friedman, Jerome H., "Greedy Function Approximation: A Gradient Boosting Machine," *The Annals of Statistics*, 2001, *29* (5), 1189–1232.

Geiser, Saul and Maria Veronica Santelices, "Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes," 2007.

Guerdan, Luke, Amanda Coston, Kenneth Holstein, and Zhiwei Steven Wu, "Predictive Performance Comparison of Decision Policies Under Confounding," in "Proceedings of the 41st International Conference on Machine Learning," Vol. 235 of *Proceedings of Machine Learning Research* 2024.

Hanley, James A. and Barbara J. McNeil, "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve," *Radiology*, 1982, *143* (1), 29–36.

Hippel, Paul T Von and Alvaro Hofflinger, "The data revolution comes to higher education: identifying students at risk of dropout in Chile," *Journal of Higher Education Policy and Management*, 2021, *43* (1), 2–23.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human decisions and machine predictions," *The quarterly journal of economics*, 2018, *133* (1), 237–293.

_ , Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer, "Prediction policy problems," *American Economic Review*, 2015, *105* (5), 491–495.

Lundberg, Scott M and Su-In Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, 2017, *30*.

Malamud, Ofer, "Discovering one's talent: learning from academic specialization," *ILR Review*, 2011, *64* (2), 375–405.

Mullainathan, Sendhil and Jann Spiess, "Machine learning: an applied econometric approach," *Journal of Economic Perspectives*, 2017, *31* (2), 87–106.

Oreopoulos, Philip and Uros Petronijevic, "Student Coaching: How Far Can Technology Go?," *Journal of Human Resources*, 2018, *53* (2), 299–329.

_ **and** _ , "The Remarkable Unresponsiveness of College Students to Nudging and What We Can Learn from It," NBER Working Paper 26059, National Bureau of Economic Research 2019.

Ost, Ben, "The role of peers and grades in determining major persistence in the sciences," *Economics of education review*, 2010, *29* (6), 923–934.

Picciano, Anthony G, "The evolution of big data and learning analytics in American higher education.," *Journal of asynchronous learning networks*, 2012, *16* (3), 9–20.

Plak, Simone, Ilja Cornelisz, Martijn Meeter, and Chris van Klaveren, "Early Warning Systems for More Effective Student Counselling in Higher Education: Evidence from a Dutch Field Experiment," *Higher Education Quarterly*, 2022, *76* (1), 131–152.

Rask, Kevin, "Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences," *Economics of education review*, 2010, *29* (6), 892–900.

Rossman, Daniel, Martin Kurzweil, and Bethany Lewis, "MAAPS Advising Experiment: Evaluation Findings after Six Years," Research Report, Ithaka S+R May 2023.

Shmueli, Galit, "To explain or to predict?," *Statistical science*, 2010, pp. 289–310.

Sorensen, Lucy C, ""Big Data" in educational administration: An application for predicting school dropout risk," *Educational Administration Quarterly*, 2019, *55* (3), 404–446.

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

Stinebrickner, Todd and Ralph Stinebrickner, "Learning about academic ability and the college dropout decision," *Journal of Labor Economics*, 2012, *30* (4), 707–748.

Tibshirani, Robert, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996, *58* (1), 267–288.

Vaarma, Matti and Hongxiu Li, "Predicting Student Dropouts with Machine Learning: An Empirical Study in Finnish Higher Education," *Technology in Society*, 2024, *76*, 102474.

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

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

A Additional Results

A.1 Prediction Specification and Validation

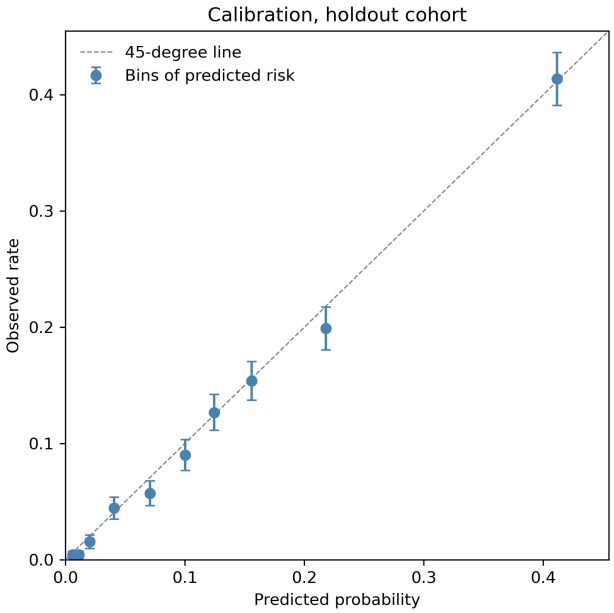


Figure A1: Calibration of XGBoost Predictions

Notes: The figure compares predicted switching probabilities from XGBoost with realized next-term switching rates in the 2015 evaluation cohort. Observations are grouped by predicted probability.

A.2 Targeting and Switching Timing

A.3 Predictive Contributions in the XGBoost Model

Table A1: Prediction Sample by Starting Major

Code	Major	Students	Student-terms
Business			
ACT	Accounting	757	4,000
BUS	Business administration	967	2,726
ECO	Economics	407	2,019
FIN	Finance	1,225	6,317
MGT	Management	365	1,326
MKG	Marketing	503	2,438
Business total		**4,224**	**18,826**
Engineering			
ARO	Aerospace engineering	458	2,616
BEG	Biological engineering	55	235
BME	Biomedical engineering	402	1,529
CHE	Chemical engineering	658	3,519
CHY	Chemistry	786	3,340
CPE	Computer engineering	402	1,935
CPS	Computer science	445	2,373
CVE	Civil engineering	241	1,568
ECE	Electrical engineering	368	2,179
EVE	Environmental engineering sciences	158	779
ISE	Industrial and systems engineering	349	2,136
IST	Information systems	76	389
MAT	Mathematics	198	977
MCE	Mechanical engineering	990	6,292
MSE	Materials science and engineering	153	774
NES	Nuclear engineering sciences	122	556
PS	Physics	156	708
STA	Statistics	56	299
Engineering total		**6,073**	**32,204**
Health			
APK	Applied physiology and kinesiology	670	3,224
BIO	Biology	2,177	10,403
CSD	Communication sciences and disorders	173	982
HEB	Health education and behavior	86	358
HES	Health science	1,392	5,864
MCB	Microbiology and cell science	631	2,949
NSG	Nursing	871	4,592
SPM	Sport management	153	844
ZLY	Zoology	155	649
Health total		**6,308**	**29,865**
Total		**16,605**	**80,895**

Counts refer to the prediction sample for the 2011 through 2015 entry cohorts. Students are classified by their starting major. Categories follow the classification used in the analysis.

Table A2: GPA Thresholds by Major and Term

Panel A: Majors with constant overall GPA thresholds

		Core-course GPA				
Major	Overall GPA	Term 1	Term 2	Term 3	Term 4	Term 5
Business						
ACT	3.00	3.00	3.00	3.00	3.00	3.00
BUS, FIN, MGT, MKG	2.00	2.50	2.75	3.00	3.00	3.00
ECO	2.00	None	2.50	2.75	3.00	3.00
Engineering						
ARO, BEG, CHE, CPE, CVE, ECE, EVE, ISE, MCE, MSE, NES	2.00	2.50	2.50	2.50	2.50	2.50
BME	2.00	3.00	3.00	3.00	3.00	3.00
CHY	2.00	None	2.50	2.65	2.75	2.75
CPS, MAT, PS, STA	2.00	None	None	None	2.50	2.50
IST	2.00	2.50	2.75	3.00	3.00	3.00
Health						
APK	2.00	3.00	3.00	3.00	3.00	3.00
BIO, MCB	2.00	2.50	2.50	2.50	2.50	2.50
HEB	2.00[a]	2.50	2.50	2.50	2.50	NA
SPM	2.00[a]	2.20	2.20	2.20	2.20	NA
ZLY	2.00	2.00	2.00	2.50	2.50	2.50

Panel B: Majors with changing overall GPA thresholds

Major	GPA measure	Term 1	Term 2	Term 3	Term 4	Term 5
CSD	Overall	2.00	2.30	2.70	2.70	2.70
	Core-course	2.00	2.30	2.70	2.70	2.70
HES	Overall	2.30	2.70	3.00	3.00	NA
	Core-course	None	None	3.00	3.00	NA
NSG	Overall	3.00	3.10	3.20	3.30	NA
	Core-course	3.00	3.10	3.20	3.30	NA

Majors with identical thresholds are grouped together within each category. Panel A reports the overall GPA threshold once because it is constant across recorded terms. Panel B reports both thresholds by term for CSD, HES, and NSG, all in the Health category. None denotes a threshold coded as zero in the data; NA indicates no record for that major and term. Core courses are the starting major's critical-tracking courses. Major names are listed in Table A1.

[a] The overall GPA threshold is recorded only for terms 1 through 4.

Table A3: Term-Ahead Switching: Full Logit Estimates

	(1) Demographics	(2) +Pre-College	(3) +Major & Term	(4) +College GPA	(5) +Course Performance	(6) +Major FE
Male	-0.0146***	-0.00351	0.000291	-0.0086***	-0.0085***	-0.00675**
	(0.00308)	(0.00331)	(0.00285)	(0.00277)	(0.00276)	(0.00278)
Underrepresented minority	0.00856**	-0.00886**	-0.00571*	-0.00433	-0.00466	-0.00646**
	(0.00355)	(0.00365)	(0.00299)	(0.00286)	(0.00285)	(0.00283)
First-generation student	0.0149***	0.00847	0.00826*	-0.00188	-0.0022	-0.00408
	(0.00553)	(0.0055)	(0.00449)	(0.00425)	(0.00424)	(0.00425)
In-state student	-0.0206***	-0.00335	-0.000383	-0.011**	-0.0104**	-0.0136***
	(0.00587)	(0.00609)	(0.00496)	(0.00483)	(0.00482)	(0.00476)
High school GPA		-0.096***	-0.0768***	-0.0148***	-0.0177***	-0.0229***
		(0.00666)	(0.00535)	(0.00533)	(0.00534)	(0.00525)
Math test score		-0.000371***	-0.000271***	-0.000119***	-0.000124***	-0.000142***
		(0.0000284)	(0.000024)	(0.0000232)	(0.0000232)	(0.0000234)
Verbal test score		0.00000692	-0.00000775	0.0000756***	0.000066***	0.0000528**
		(0.0000295)	(0.0000242)	(0.0000232)	(0.0000232)	(0.0000231)
Major credit requirement			-0.000262	-0.00198***	-0.00198***	
			(0.000428)	(0.000425)	(0.000425)	
STEM major			-0.0484***	-0.0453***	-0.0458***	
			(0.00403)	(0.00398)	(0.00395)	
Major median earnings, 10 years			0.000000148*	0.000000258***	0.00000022***	
			(8.17e-08)	(7.79e-08)	(7.87e-08)	
Cumulative GPA				-0.107***	-0.0904***	-0.086***
				(0.00322)	(0.00397)	(0.00401)
Cumulative core-course GPA				-0.0109***	-0.0107***	-0.0139***
				(0.0015)	(0.00148)	(0.00151)
Cumulative course easiness				0.0793***	0.0686***	0.0821***
				(0.0107)	(0.0108)	(0.0119)
Change in term GPA from last term					-0.000613	0.000325
					(0.00358)	(0.00355)
Change in term easiness from last term					-0.0208**	-0.0205**
					(0.00867)	(0.00859)
Number of D/F grades this term					0.0235***	0.0239***
					(0.00339)	(0.00338)
Credits this term					-0.0012***	-0.00161***
					(0.000352)	(0.00035)
Term indicators	No	No	Yes	Yes	Yes	Yes
Major fixed effects	No	No	No	No	No	Yes
Observations	62906	62906	62896	62896	62896	62896

Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Switch Timing and Cumulative Cost at Graduation

	Credits	Credits	Terms	Terms
Term 3	0.745	1.197**	0.034	0.093**
	(0.600)	(0.580)	(0.039)	(0.038)
Term 4	2.571***	3.463***	0.146***	0.241***
	(0.649)	(0.634)	(0.042)	(0.041)
Term 5	3.471***	4.362***	0.490***	0.512***
	(1.141)	(1.048)	(0.087)	(0.080)
Term 6	6.475***	7.761***	0.640***	0.672***
	(1.451)	(1.404)	(0.103)	(0.090)
Term 7 or later	20.326***	18.991***	1.768***	1.638***
	(2.571)	(2.643)	(0.154)	(0.140)
Demographic controls	No	Yes	No	Yes
Academic controls	No	Yes	No	Yes
Major FE	No	Yes	No	Yes
Observations	5161	5161	5161	5161
R^2	0.021	0.143	0.039	0.175

Standard errors in parentheses.

Sample: graduates with an observed switch term, N = 5,161. Each row is the last term a student spends in the starting major; the switch is observed the following term. The omitted category is term 2. Group sizes for terms 2 through 7 or later are 2,379, 1,271, 1,043, 262, 145, and 61.

Outcomes are cumulative credits at the last enrolled term and terms to degree. Demographic controls are sex, underrepresented minority status, first-generation status, and in-state status; academic controls are cumulative GPA, first-year D or F courses, and first-year credits; controlled columns also include entry-year fixed effects.

Descriptive, not causal: graduates only, so late switchers who left the university are not observed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Targeting Performance by Term

	Contacts K_t	Current system	XGBoost	Logit	Random
Term 1	969	230	292	283	167
Term 2	491	233	259	239	107
Term 3	336	108	126	111	42
Term 4	250	85	93	70	30
Term 5	168	30	41	28	7
Total	2,214	686	811	731	354

Notes: The table uses the 2015 evaluation cohort and restricts the risk set in each term to students who have not yet left their starting major. Entries report contacted student-terms followed by a switch in the next term. In each term, XGBoost, logit, and random selection use the same contact capacity, K_t, as the university's current system. Random reports the expected number of subsequent switchers reached at the same contact rate.

Table A6: Who Appears at the Top of the XGBoost Ranking?

	Top 10 percent	Remaining 90 percent
Core-course GPA	2.23	3.32
Overall GPA	2.99	3.57
D or F courses	0.32	0.02
Term number	2.10	4.16
STEM starting major	0.74	0.93
Actual next-term switch	0.414	0.077

Notes: The unit of observation is the student-term. Groups are defined by the XGBoost predicted probability in the 2015 evaluation cohort. Entries are group means. D or F courses is the number of D or F grades received in the current term.

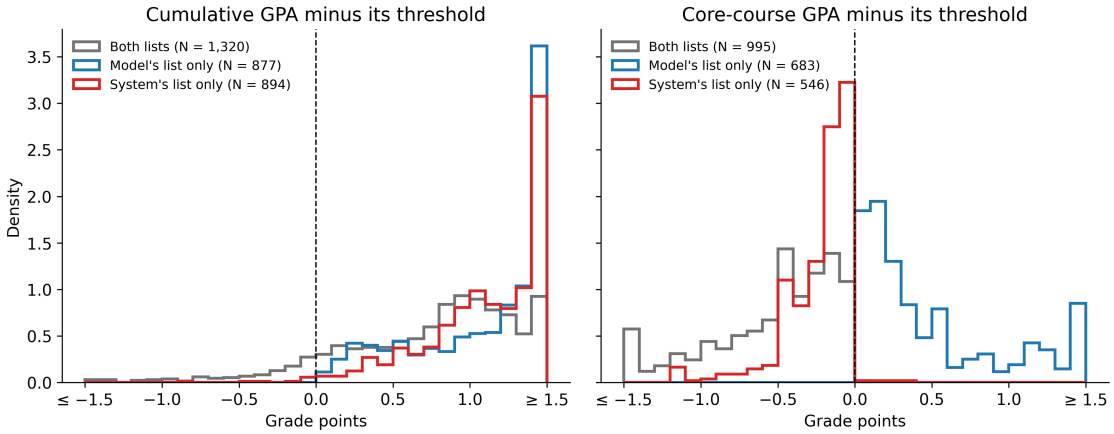


Figure A2: GPA Relative to Advising Thresholds

Notes: The figure compares students selected by both methods, only the current system, and only XGBoost in the 2015 evaluation cohort. GPA is measured relative to the corresponding threshold. The left panel includes observations with an applicable overall GPA threshold. The right panel includes observations with a positive core-course GPA threshold and completed core courses.

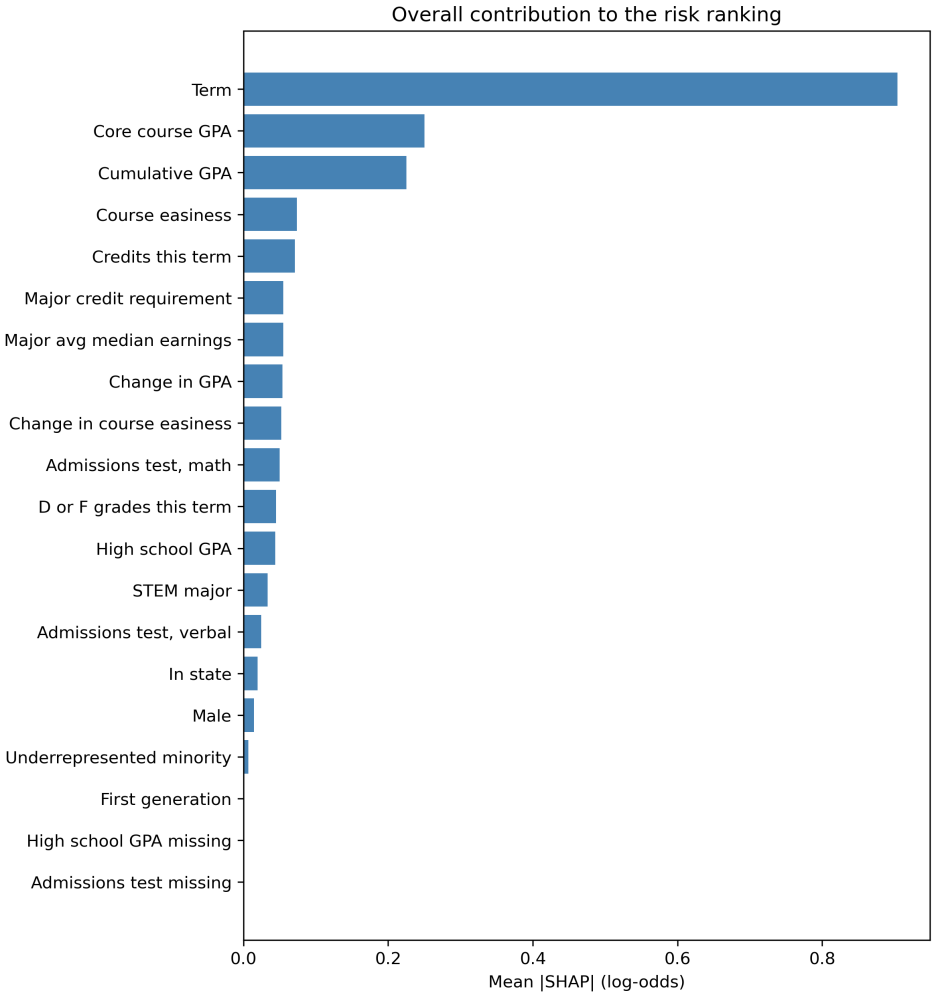


Figure A3: Mean Absolute SHAP Values for the XGBoost Model

Notes: Bars report mean absolute SHAP values across student-term observations in the 2015 evaluation cohort. SHAP values are measured in log-odds. Larger values indicate a greater average contribution to the fitted prediction, regardless of whether the contribution raises or lowers predicted switching risk. Starting-major indicators remain in the fitted model and are omitted from the displayed ranking for readability.

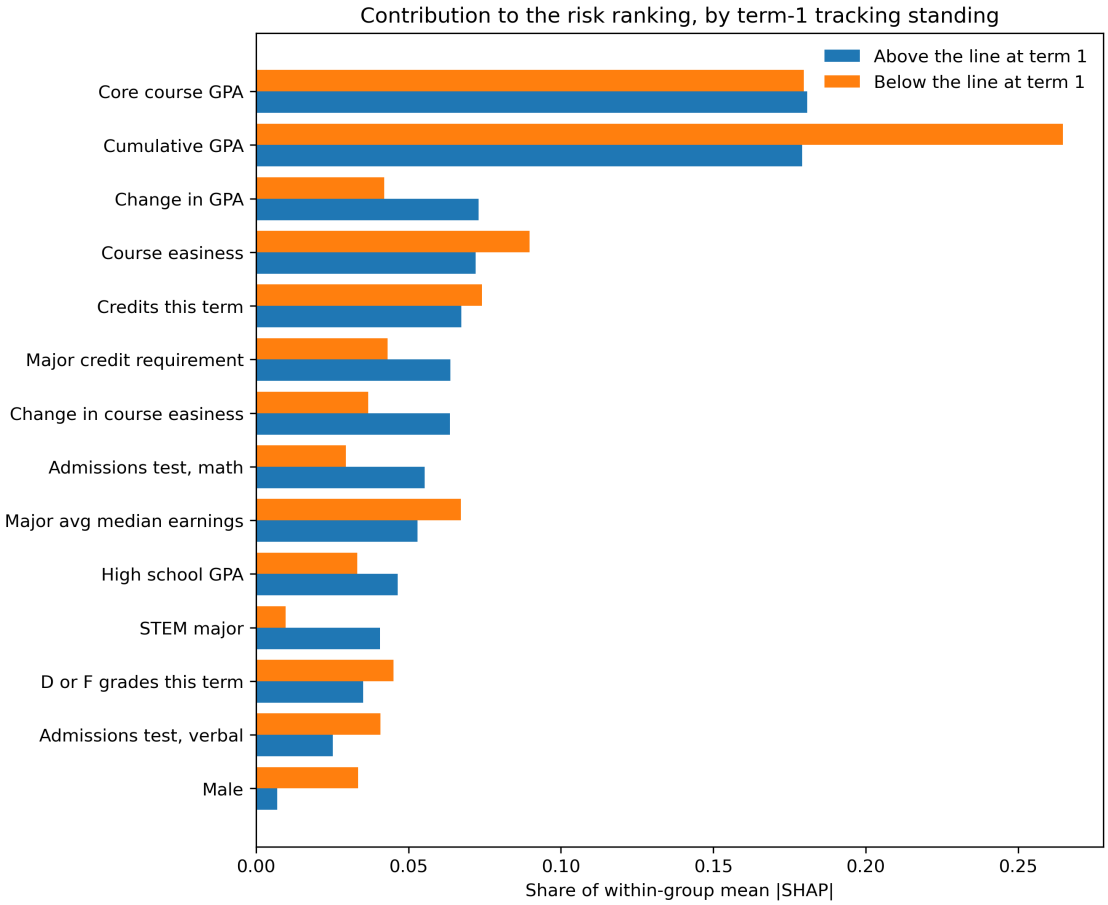


Figure A4: Predictive Contributions by Term-1 Tracking Standing

Notes: Each bar reports a predictor's mean absolute SHAP value as a share of the total across predictors within the group. Students are grouped by whether they were below a tracking threshold in term 1. This classification is fixed for the student's subsequent observations, and students without a term-1 standing are excluded. One model is fitted separately for each group using student-term observations from the development cohorts and the hyperparameters of the pooled model. SHAP values are computed for the corresponding group in the 2015 evaluation cohort. Starting-major indicators are omitted from the figure. Predictors are shown if their share is at least 0.03 in either group.

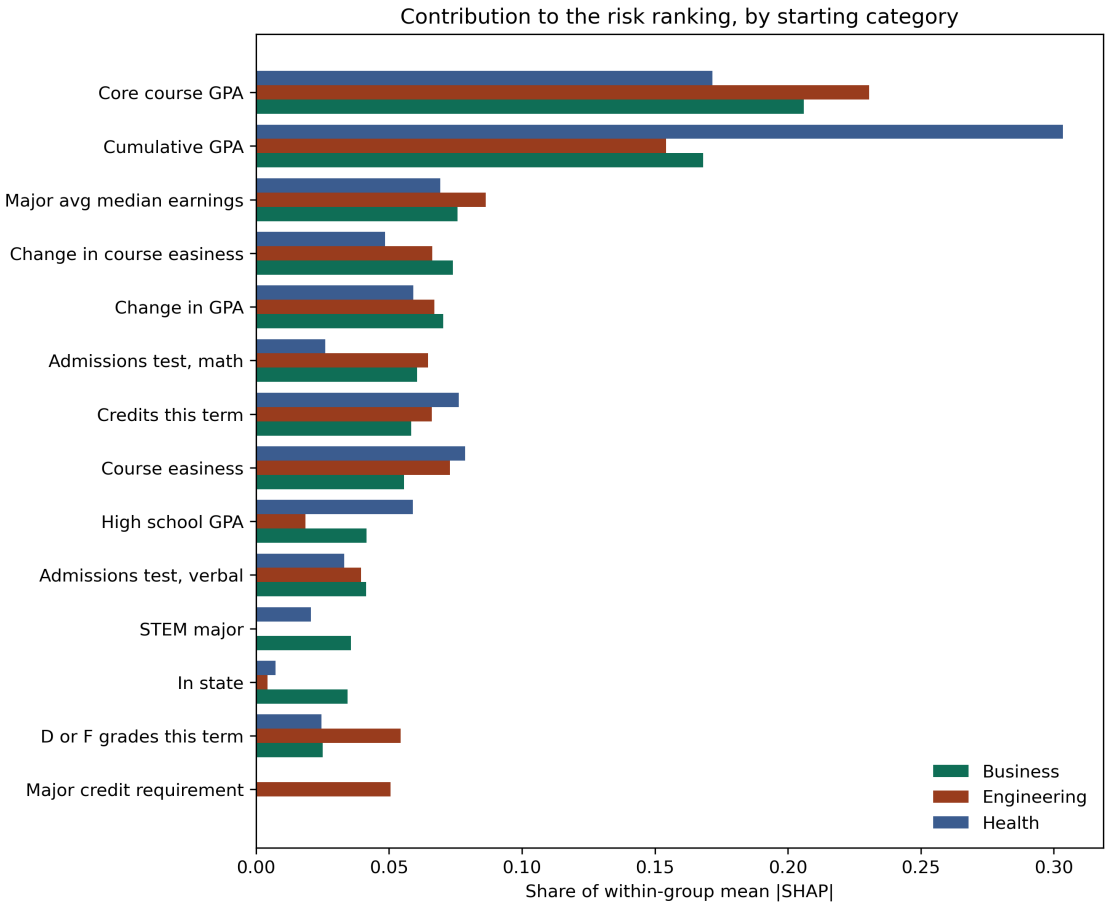


Figure A5: Predictive Contributions by Starting Category

Notes: Each bar reports a predictor's mean absolute SHAP value as a share of the total across predictors within a starting-major category. One model is fitted separately for each category using student-term observations from the development cohorts and the hyperparameters of the pooled model. SHAP values are computed for the corresponding category in the 2015 evaluation cohort. Starting-major indicators are omitted from the figure. A major attribute with no within-category variation contributes nothing by construction, so shares for major attributes are not directly comparable across categories. Predictors are shown if their share is at least 0.03 in any category.