Openness to AI in Higher Education*

Fernando Saltiel‡

This version: September 29, 2026
First version: March 26, 2026

Please click here for the latest version.

Abstract

This paper examines how AI is being integrated in higher education. I construct a dataset of 327,000 syllabi from 50 public universities between 2024 and 2026, and use open-weight large language models to classify how instructors are incorporating AI into the classroom. While only 7 percent of courses integrate AI into teaching, this share has more than doubled in the past two years. Moreover, this integration is concentrated at selective, research-intensive universities, while less selective institutions rely more on explicit policies on AI use. Instructors play a key role in driving AI adoption, as classroom integration more than doubles when a professor with prior AI teaching experience takes over a course. Assistant professors and research-active instructors are especially likely to integrate AI into their courses, particularly at selective universities.

*I thank Maggie Jones, Fabian Lange, Matias Cortes, Brennan McLachlan and Jérôme Larivière for detailed comments and suggestions, as well as seminar participants at the NBER Summer Institute Digital Economics and Artificial Intelligence, the Barcelona Summer Forum and McGill University for helpful feedback. Magnolia Miller and Aditi Ramanand provided excellent research assistance. All errors are my own.

†A companion website is available at: https://ai-higher-ed.netlify.app.

‡McGill University and IZA. Email: **fernando.saltiel@mcgill.ca**

1 Introduction

The emergence of generative artificial intelligence tools has prompted rapid adoption across the economy, with recent evidence documenting sizable effects on worker productivity, task allocation, and organizational design (Brynjolfsson et al., 2025; Bick et al., 2025). The diffusion of AI creates distinct challenges for universities, as they must decide how to integrate these tools into the classroom even as students are already using them in their coursework (Contractor and Reyes, 2025; Handa et al., 2025). These decisions matter not only because universities regulate AI use in the classroom, but also because they shape the skills that students develop in college. Understanding how universities and instructors are integrating AI in the classroom is therefore increasingly important.

In this paper, I analyze how artificial intelligence is being integrated into higher education. In particular, I examine how instructors are incorporating AI into their classes by constructing a dataset of course syllabi across public universities through Fall 2026. Syllabi are a natural context for studying this question, as they capture the content of college classes and the approaches that professors take in the classroom (Biasi and Ma, 2022). As such, they reveal how instructors are actually integrating AI into their classes, ranging from what they teach to how they evaluate student work.

To this end, I collect publicly available course syllabi, recovering 327,495 syllabi across 50 public universities in seven states between 2024 and 2026, covering a period with widespread AI use in higher education.[1] To understand which professors are most likely to integrate AI into the classroom, I construct a complementary dataset capturing their research productivity and academic rank from multiple data sources. I also recover institutional and major-level characteristics that allow me to identify which universities and majors are driving this integration.

The share of syllabi mentioning AI has risen sharply, from 43 percent in Fall 2024 to 75 percent by Fall 2026, up from fewer than three percent before ChatGPT. However, a syllabus that prohibits AI use and one that builds assignments around it would both be classified as mentioning the technology, despite taking opposite approaches in the classroom. I instead aim to recover how instructors are integrating AI into their teaching and how they use it to evaluate their students.

I thus construct four measures that capture how AI is being integrated into the classroom. On the teaching margin, I measure whether AI is part of what the course teaches, through its class sessions, topics, or course materials. On the evaluation margin, I measure whether at least one graded task requires students to engage with AI. I also consider whether the syllabus includes an explicit policy around the use of AI, and whether it affirmatively permits students to use it. Lastly, I validate these measures with human raters, finding high inter-rater agreement across all four questions (Dell, 2025; Haaland et al., 2025).

I classify whether these measures are present in each syllabus using large language models (LLMs), which are well-suited for this task as they can read unstructured syllabus

[1] For eight of these universities, I use syllabi dating back to 2019, which allows me to analyze how AI integration in the classroom has evolved since the release of ChatGPT.

text in context (Ash and Hansen, 2023). I implement the classification approach using local open-weight models, which enhance the reproducibility of the results (Korinek, 2023).[2] Specifically, I classify each syllabus using three independent open-weight models, taking the modal answer across them. To ensure that this approach can reliably classify the integration measures at scale, I compare the LLM classifications against the human raters. The LLM-based approach passes the Alternative Annotator Test, as it matches the consensus of human raters as often as the individual raters do (Calderon et al., 2025).

I use the LLM-based classification to assess how AI has been integrated into higher education. I find that 73 percent of courses had explicit AI policies by Fall 2026. However, just 37 percent of syllabi actually allowed students to use AI. This suggests that instructors are reticent to allow this technology in the classroom. In fact, just six percent of courses incorporate AI into their teaching. Professors are even less likely to incorporate AI into student evaluation, as only three percent of classes include assignments requiring students to engage with this technology. Altogether, a small share of students are being required to actively engage with AI in their college classes.

At the same time, AI integration could vary across universities, given differences in resources, research productivity and student populations. Eight percent of courses at selective institutions have already integrated AI into the classroom, compared to just five percent at less selective universities. Moreover, R1 institutions and those with higher research expenditures and faculty salaries are more likely to have integrated AI into their classes. This evidence is thus consistent with prior work showing that AI adoption is concentrated at larger and more research-intensive firms (McElheran et al., 2024; Babina et al., 2024). Within these universities, graduate courses integrate AI at higher rates than undergraduate classes, consistent with the fact that they cover content closer to the research frontier (Biasi and Ma, 2026).

Since majors differ significantly in their exposure to AI (Eloundou et al., 2024), I also analyze how adoption varies across fields. Business majors are most likely to have incorporated AI into their classes, including 20 percent of courses in Marketing. By contrast, courses in the Social Sciences have the lowest rates of classroom adoption: just two percent of Economics courses have integrated AI. Importantly, the majors most exposed to AI show the highest rates of classroom adoption. Instructors in these fields may be adapting their courses to a labor market increasingly shaped by AI (Bick et al., 2025).

Given the large differences in AI integration across universities and majors, I estimate a variance decomposition that separates the contributions of instructors, courses, and institutions. Professors are the main drivers of whether AI is adopted in the classroom, as they account for 25 percent of the variation across the teaching and evaluation margins. Moreover, course-level differences account for a comparable share of the variation in these outcomes. On the other hand, universities and majors play a limited role in shaping whether AI is integrated into the classroom.

Since professors are the main drivers of whether AI is adopted in the classroom, I ex-

[2]For each syllabus, the LLMs classify each of the four questions independently, to ensure that the response to one question does not influence its response to the others.

amine which instructors lead this adoption. Assistant professors are more likely than full professors to have incorporated AI in their classes, conditional on the courses they teach and the broader institutional environment. Research activity also predicts adoption outcomes, as instructors who published recently are more likely to integrate AI in their classes. These effects are largely for integration on the teaching margin, and are robust to additional controls and fixed effects. Taken together, the leading edge of classroom AI integration is concentrated among young, research-active faculty members.

These differences hold in STEM and non-STEM fields alike. The career-stage differences, however, are largest at selective universities, where assistant professors are especially likely to integrate AI into the classroom. At the same time, publication and citation counts do not predict AI integration among instructors teaching the same course. This differs from Biasi and Ma (2022), who show that more productive researchers are more likely to bring frontier knowledge into the classroom. AI integration follows a different logic, since it requires instructors to experiment with a new technology and to redesign their courses around it. That helps explain why young, research-active faculty members are more likely to integrate AI into the classroom.

This paper makes several contributions to the literature. First, I contribute to a rapidly growing literature documenting the adoption of artificial intelligence across firms, workers, and occupations (Acemoglu et al., 2022; McElheran et al., 2024; Bonney et al., 2024; Brynjolfsson et al., 2025; Bick et al., 2025; Humlum and Vestergaard, 2025; Dillon et al., 2025; Hartley et al., 2025). I contribute to this literature by showing how AI is being adopted in higher education. I do so by developing an approach that allows me to recover how professors are integrating AI into the classroom from unstructured syllabus data. This allows me to uncover multiple margins of integration, and to distinguish actual classroom adoption from broader AI policy language (Chirikov, 2026). As such, I also contribute to recent work on how students are using artificial intelligence in higher education (Contractor and Reyes, 2025; Handa et al., 2025).

I further contribute to a literature examining how instructors shape the content of what is taught in higher education. Most closely related to this paper is Biasi and Ma (2022), who develop the approach of using course syllabi to recover curricular content at scale and show that research-active instructors are more likely to teach frontier knowledge in their courses. I contribute to this work by introducing an approach that uses large language models to reliably measure how instructors integrate AI into the classroom. Moreover, I show that AI integration has been led by young, research-active faculty members, a pattern that holds across fields and is most pronounced at selective universities.[3] I also find that professors are the main drivers of AI integration in the classroom, fitting in with a broader literature showing how instructors shape outcomes in higher education (Hoffmann and Oreopoulos, 2009; Carrell and West, 2010; Braga et al., 2016; De Vlieger et al., 2016; Feld et al., 2020).

Lastly, I contribute to a growing literature using large language models to recover nuanced information from text across different contexts (Bartik et al., 2024; Link et al., 2024;

[3]This aligns with recent evidence showing AI adoption is concentrated among younger workers (Humlum and Vestergaard, 2025; Bick et al., 2025).

Braghieri et al., 2024; Bursztyn et al., 2025; Lagakos et al., 2025; Clayton et al., 2025; Chyn et al., 2025; Ash et al., 2025). Here, I show that large language models can reliably recover measures that require interpretation from unstructured syllabus text. I implement the approach using multiple local open-weight models, which makes the measurement approach transparent and reproducible (Korinek, 2023; Ludwig et al., 2025). I also validate the resulting measures against independent human raters to show that these models can reliably measure AI integration in higher education (Dell, 2025; Haaland et al., 2025). This paper thus also fits into a broader literature using text-as-data methods in economics (Baker et al., 2016; Gentzkow et al., 2019; Ash and Hansen, 2023), as I recover rich information from course syllabi on how instructors adapt their classes to a new technology.

The rest of the paper is structured as follows. Section 2 describes the data sources and presents descriptive statistics on the sample of university syllabi. Section 3 introduces the classification framework and measurement approach. Section 4 documents the prevalence and evolution of AI integration, along with its heterogeneity across majors, universities and courses. Section 5 studies the drivers of AI integration, including the variance decomposition and instructor changes. Section 6 examines who integrates AI into the classroom. Section 7 concludes.

2 Data Sources and Descriptive Statistics

To understand how AI has been integrated into higher education, I take advantage of course syllabi across public universities in the United States. College syllabi typically include detailed information on course content, topics covered, specific assignments and course policies, which allows me to observe how instructors incorporate AI into their classes. Following the rapid rollout of generative AI tools in higher education, I assemble a new dataset of publicly available syllabi from 50 public universities between 2024 and 2026. Moreover, I draw on a variety of external data sources to recover detailed characteristics of instructors, universities, college majors, and courses, to examine which characteristics drive AI adoption in the classroom.

2.1 University Syllabi

I construct this dataset by using publicly available syllabi from public universities across seven states.[4] I recover course syllabi for 50 universities, covering selective research universities such as UT Austin and the University of Georgia, as well as broad-access institutions such as Florida A&M and West Texas A&M. For each syllabus, I recover the main information for the course, drawing on both structured fields in each university's data system and the syllabus text itself.

The main estimation sample includes syllabi for the courses offered between Fall 2024 and Fall 2026. I focus on these five academic terms, which cover the period of widespread AI availability in higher education. Each syllabus in the main sample corresponds to a

[4] Appendix B.1 describes the data sources and the construction of the main sample.

course taught by a specific instructor in a given semester. As such, the main sample includes 327,495 syllabi covering 107,705 courses taught by 78,383 different instructors in the 50 universities in the main sample (Table A1). For eight of these universities, I additionally recover a longitudinal sample of 120,179 syllabi between Fall 2019 and Spring 2024, which I use to analyze outcomes before and after the rollout of ChatGPT.

Beyond the main course characteristics recovered from the syllabi, I extract a set of text-based features from each syllabus that capture relevant information about the course and the instructor (Appendix B.3). As a first pass at understanding how AI is being integrated into the classroom, I identify syllabi that include AI-related terms such as artificial intelligence, generative AI, large language model, and ChatGPT (Appendix B.3.6). This allows me to provide an initial description of the general prevalence of AI-related content across courses.

2.2 Instructor Characteristics

To examine which instructors are incorporating AI into their courses, I complement the syllabus data with a rich set of instructor characteristics. I draw on publicly available salary records, bibliometric databases, instructors' CVs, faculty directories, and institutional profiles to construct a detailed instructor-level dataset. These data allow me to examine how academic rank, gender, research productivity, and career experience drive AI adoption in the classroom.

To measure instructors' research productivity, I match the instructor roster to OpenAlex, an open-access bibliometric database (Priem et al., 2022). I match 47 percent of instructors in the sample (Appendix B.2), with cross-university variation in match rates reflecting institutional differences in faculty composition. From the matched profiles, I recover publication counts and dates, citation counts, and h-index as measures of research productivity and impact. I also recover information on instructors' educational backgrounds from ORCID education records, faculty CVs, and departmental profile pages. Altogether, these data sources allow me to identify whether an instructor has any publications, along with their PhD institution and completion year.

I also recover academic rank from state salary databases, which record instructors' official institutional titles. I classify each instructor as a full, associate, or assistant professor, or as non-tenure-track faculty, which includes lecturers, adjuncts, and instructors of practice. For instructors' gender, I use direct records from the Texas salary database, along with a name-based probabilistic assignment using Social Security Administration baby-name data. All in all, I recover gender and academic rank for 90 and 72 percent of instructors in the sample, respectively (Table A2).

2.3 University, Major, and Course Characteristics

To understand the extent to which the integration of AI into the classroom varies across universities, I also assemble a set of institutional characteristics from multiple data sources (Appendix B.1). First, I use data from the Integrated Postsecondary Education Data System

(IPEDS) to capture measures of research intensity, faculty compensation, and average SAT scores. I additionally recover the median parental income for each institution from the Mobility Report Cards (Chetty et al., 2020). I further classify the 50 universities into three selectivity tiers based on the average SAT scores of their entering students, ranging from selective research universities to broad-access institutions.

I similarly recover a set of major-level characteristics from multiple data sources. In particular, I capture the mid-career median earnings, unemployment rate, gender composition, and share of graduates who earn a graduate degree for each major from the Federal Reserve Bank of New York's College Labor Market data (Abel et al., 2014). I also construct a measure of average SAT scores in each major from the National Postsecondary Student Aid Study (Appendix B.3). To understand the extent to which different majors are exposed to AI, I use the exposure measure proposed by Felten et al. (2023), which maps AI exposure scores from occupations to college majors. This measure intuitively captures the fact that AI exposure is higher for English and Marketing majors than for those in Nursing.

Since the 50 universities in the main sample use 7,472 distinct department codes, I harmonize them into 42 majors. For instance, I classify courses from Neuroscience, Microbiology, and Biochemistry departments into a harmonized Biology major (Table B3). I further aggregate the 42 majors into five macro-fields, covering STEM, Humanities, Social Sciences, Business, and Other. The harmonized majors show significant heterogeneity in AI exposure, earnings, and SAT scores across fields.

I also recover the academic level of each course from its university's course numbering conventions. Since nearly all universities follow a standard system in which the leading digit of the course number indicates the year of study, I classify courses as freshman, sophomore, junior, senior, or graduate (Appendix B.3). I use this five-level classification as the main variable, and also classify courses as basic undergraduate, advanced undergraduate, or graduate following Biasi and Ma (2022).[5]

2.4 Descriptive Analysis

Table 1 presents descriptive statistics for the main sample, pooling the five terms from Fall 2024 through Fall 2026 across the 50 universities. Half of the courses in the sample are taught by tenure-track faculty, while 53 percent are taught by instructors who have published at least one paper, consistent with the composition of the instructor workforce at these institutions.[6]

The second column shows the share of syllabi that mention artificial intelligence, which 64 percent of syllabi in the sample do. Moreover, this share rises from 46 percent in the 2024–25 academic year to 75 percent in Fall 2026. Women are more likely to mention AI in their course syllabi than men, but the corresponding differences across professors' tenure-track status are smaller. Lastly, undergraduate courses are more likely to mention AI than graduate courses, and AI mentions are also more common at broad-access universities.

[5] I also use the Florida Statewide Course Numbering System, which provides harmonized course identifiers across Florida universities and allows me to compare syllabi for the same course across institutions.

[6] Table A3 shows the corresponding evidence for the syllabi in the longitudinal sample prior to Fall 2024.

Table 1: Descriptive Statistics

	Share of Sample (1)	**Share Mentioning AI** (2)
Instructor Gender		
Female	47.2	68.5
Male	52.8	60.5
Instructor Characteristics		
Tenure-Track	50.3	62.7
Any Publications	53.4	62.4
University Characteristics		
Selective	32.1	55.8
Moderately Selective	39.7	64.0
Broad Access	28.2	73.1
Major Characteristics		
Business	9.6	65.2
STEM	39.3	60.3
Social Sciences	20.7	67.6
Humanities	28.4	65.7
Course Characteristics		
Undergraduate	78.6	64.8
Graduate	21.4	61.0
Academic Year		
2024–25	24.8	46.2
2025–26	42.6	65.8
2026–27	32.6	75.1
Observations	327,495	209,243

Note: Table 1 presents descriptive statistics for the main sample of syllabi, pooling the five terms from Fall 2024 through Fall 2026. Column (1) shows the share of syllabi in each row, and column (2) shows the share of syllabi in each row that mention artificial intelligence. For the gender and tenure-track rows, I focus on syllabi with non-missing instructor information. University selectivity follows the three tiers based on the average SAT scores of entering students. Business, STEM, Social Sciences and Humanities denote the harmonized macro-fields, and Undergraduate combines freshman through senior courses. The 2026–27 academic year includes only the syllabi from Fall 2026. The last row reports the number of syllabi in the sample and the number that mention artificial intelligence.

Figure A1 shows there is significant heterogeneity in the prevalence of AI mentions across the harmonized majors in the sample. For instance, 72 percent of syllabi in Sociology mention AI, compared to 57 percent in Economics. Meanwhile, course syllabi in STEM majors are the least likely to include explicit references to artificial intelligence.

I draw on the longitudinal sample to measure the evolution of AI mentions in course syllabi dating back to 2019–20. A small share of syllabi already discussed artificial intelligence prior to the introduction of ChatGPT, particularly in Information Systems, Computer Science and Marketing courses. While the overall prevalence of AI mentions in course syllabi was substantially lower in this period, it increased substantially following the rollout of ChatGPT in November 2022 (Figure A2). In particular, the first significant response emerged in the Fall 2023 semester at seven of the eight universities. By Fall 2026, 87 percent

of syllabi at UT Austin and 48 percent at UT Dallas mention artificial intelligence.

While this analysis shows that a significant share of syllabi now mention artificial intelligence, this approach combines fundamentally different approaches into a single measure. For instance, a syllabus stating that "AI-generated submissions constitute academic dishonesty" and one stating that "students are expected to use ChatGPT for the course project" both mention AI, but they capture diametrically opposed pedagogical stances.

3 Measuring AI Integration in the Classroom

In this section, I introduce an approach that captures how instructors are integrating artificial intelligence into higher education. I focus on instructors as they directly shape what students learn in college, and their decision to integrate AI into the classroom may shape students' familiarity with these tools. Integrating AI into a course is not the same as keeping course content up to date with recent advances in the field, which flows naturally from a professor's own research activity. AI integration instead requires deliberate pedagogical redesign around a general-purpose technology that cuts across fields (Bresnahan and Trajtenberg, 1995).

To understand what instructors are actually doing with AI in their classrooms, I develop measures that capture how they are integrating it into their teaching and evaluation. I focus on the teaching margin, as a professor who schedules a session on prompt engineering or designs a hands-on AI exercise is committing class time to engaging with artificial intelligence. A typical undergraduate semester includes 14-16 weeks of instructional time, so these decisions displace traditional content and reveal the instructor's judgment about what knowledge students need (Comin and Hobijn, 2010). Similarly, I analyze how professors integrate AI into student evaluation, as they may have also created and adapted their assignments in response to this new technology (Acemoglu and Restrepo, 2019).

3.1 AI Integration in Teaching and Evaluation

To understand the prevalence of AI classroom integration, I first worked with research assistants who read a large number of syllabi to identify the main ways in which professors incorporate AI into their teaching and evaluation (Haaland et al., 2025). Based on this analysis, I construct a teaching integration measure that captures whether AI is part of what the course teaches, through its class sessions, topics, or course materials. In particular, this measure identifies courses that dedicate a session or module to an AI topic, describe a class activity or demonstration involving AI, or assign readings or other materials specifically about AI, signaling that the instructor views AI literacy as knowledge worth developing (Javadian Sabet et al., 2024). This margin is central to understanding classroom adoption, as it reflects a deliberate change to the content that students are exposed to.

On the evaluation margin, I consider whether at least one graded or assessed task requires students to engage with AI. This measure identifies assignments in which students must use an AI tool, submit a deliverable that includes or builds on AI output, or complete

a required step that has them produce, compare, critique, or revise AI content. This margin captures instructors who bring AI directly into graded work, whether to help students develop skills in using these tools or simply to allow AI-assisted work in their assignments. Altogether, these measures capture how instructors have redesigned their courses to explicitly integrate artificial intelligence into their teaching and evaluation.

Beyond the classroom integration margin, I also examine whether professors explicitly address AI in their syllabi and what stance they take toward its use. In this context, I analyze whether the syllabus contains any explicit policy regarding student AI use, and if so, whether the instructor explicitly allows students to use artificial intelligence in their course. These two measures capture the difference between instructors who address AI and those who affirmatively welcome it into their courses.

To capture these dimensions from the syllabi, I construct the measures as binary questions, as this format tends to yield high inter-rater reliability rates (Dolnicar et al., 2011). Each question includes brief guidance clarifying the classification criteria, following standard practice in content analysis (Lombard et al., 2002; O'Connor and Joffe, 2020). Table 2 presents the four classification questions with their accompanying guidance.

Table 2: AI Integration Classification Questions and Guidance

AI Policy. Does the syllabus contain an explicit policy, rule, or guideline about student use of AI or generative AI tools (such as ChatGPT, Copilot, or similar tools) for coursework?
Guidance: True if the syllabus states rules, permissions, prohibitions, guidelines, or expectations about whether and how students may use AI tools for coursework. The policy may come from the instructor or the university, may appear anywhere in the syllabus, and may permit, prohibit, restrict, or conditionally allow AI use.
AI Permitted. Does the syllabus explicitly state that students are permitted, allowed, or welcome to use AI tools for coursework in this course?
Guidance: True if the syllabus contains an explicit statement permitting, allowing, welcoming, or approving student use of AI tools for coursework.
Teaching Integration. Is AI part of what this course teaches: its class sessions, topics, or course materials?
Guidance: True if the course dedicates class time or course content to AI: a schedule, module, or session entry on an AI topic, a class activity, lecture, discussion, or demonstration involving AI, or an assigned reading, video, tutorial, or other course material.
Evaluation Integration. Does at least one graded or assessed task require students to engage with AI?
Guidance: True if a specific assignment, project, exam, or task requires AI in the student's own work: the task requires students to use an AI tool, the deliverable must include or build on AI output, or a required step has students produce, compare, critique, or revise AI content.

Note: Table 2 presents the four binary classification questions used to measure AI integration in course syllabi. I classify each measure independently on the full syllabus text. Appendix C presents the full coding guidance with worked examples for each question.

Before applying these measures at scale, I validate these constructs against human raters. In particular, four students independently classified a sample of 60 syllabi across all four measures.[7] The four raters agree on 95 percent of classifications on average across the four

[7] All four raters classify 20 core syllabi, and the remaining 40 are rotated so that each rater classifies 50 in

questions (Appendix C). While agreement rates are higher for questions with clear textual markers than for those requiring more interpretive judgment, these results provide a reliable benchmark for the LLM-based classification (Dell, 2025; Haaland et al., 2025).

3.2 LLM Classification Implementation

While the human raters described in Section 3.1 agree on 95 percent of classifications, coding 327,495 syllabi by hand would be prohibitively time-consuming. To classify these syllabi at scale, I take advantage of large language models, which process text by attending to the full context of each word, resolving ambiguities that keyword matching cannot (Ash and Hansen, 2023; Dell, 2025).

I implement the classification approach using open-weight models with fixed, published parameters, which I run locally to ensure the results are reproducible (Korinek, 2023). This avoids well-known problems with proprietary models, which are frequently updated or deprecated, and do not yield reproducible results even with fixed temperatures (Barrie et al., 2024). I classify each syllabus using three independent open-weight models and take the modal answer across them. I validate the resulting classifications against human raters, following the recommendation in Ludwig et al. (2025) to benchmark automated classifications against human judgments for the specific task at hand.

For each syllabus, I submit independent calls for each of the four classification questions, so that a model's response to one question does not shape its evaluation of the remaining questions. In addition to requiring a binary classification, I ask each model to provide a confidence rating, a brief justification, and direct evidence from the syllabus text when the answer is affirmative. This allows me to assess the quality of the model's reasoning and to recover detailed information on the ways in which AI appears in the classroom. In particular, I classify each question using three open-weight models, Gemma-4-31B, Qwen3.8-27B and Muse-Glimmer-30B, which rank among the top-performing open-weight models on standard evaluation leaderboards (Chiang et al., 2024).[8] As such, each model independently classifies all syllabi, providing three complete sets of classifications.

For each question, I define the final classification as the modal answer across the three models. The evidence quotes that feed the qualitative analysis of how instructors integrate AI into the classroom come from a single default rater.

The three models generally agree in their answers across the two integration questions. Pairwise agreement rates range from 98 to 99 percent, highest for the evaluation question (Table C4).

3.3 Validation

I validate the LLM-based classification approach using the Alternative Annotator Test (AAT) of Calderon et al. (2025), which evaluates whether an automated classifier can serve

total, with stratification by AI mention status, institution, and classification profile (Krippendorff, 2004).

[8] I set the temperature, presence penalty, and frequency penalty parameters to zero, the top_p parameter and the repetition penalty to one, and fix the random seed, to maximize deterministic outputs (Appendix C).

as a credible substitute for human coders. At its core, the test asks whether the automated classifier agrees with the human consensus at least as often as the individual human raters do. The logic of this test is as follows: for each syllabus-question pair, I hold out one of the four human raters and compute the modal response of the remaining three as the benchmark. I then compare whether the modal LLM label or the held-out rater is more likely to agree with this benchmark. The LLM passes the test if it matches the consensus at least as well as the held-out rater, indicating that it can be reliably used to extend the human classifications to the full set of syllabi.

The average leave-one-out agreement rate among the four human raters equals 97 percent, and the modal LLM label agrees with the modal human response in 95 percent of syllabus-question pairs. The proposed classification approach thus comfortably passes the Alternative Annotator Test, as it performs as well as humans in 96 percent of questions, far exceeding the required 50 percent threshold.

Moreover, each individual model performs similarly well against the human modal response (Appendix C). Altogether, these validation results establish that the LLM-based classification reliably measures AI policies and integration at scale, achieving agreement with human raters that is comparable to the agreement among the raters themselves. I therefore use the modal classification to characterize AI integration across the full sample of syllabi.

4 AI Integration in Higher Education

4.1 Prevalence of AI Integration

Figure 1 presents the extent to which artificial intelligence has been integrated into higher education in the Fall 2026 semester. First, close to 73 percent of courses in the sample include an explicit AI policy in their syllabi.[9] This is largely in line with the share of syllabi that explicitly mention AI, as discussed in Table 1. At the same time, just 37 percent of courses in the Fall of 2026 allow students to use these tools for their classes. The majority of professors thus have AI policies in place but do not explicitly endorse students using AI in their classes.[10]

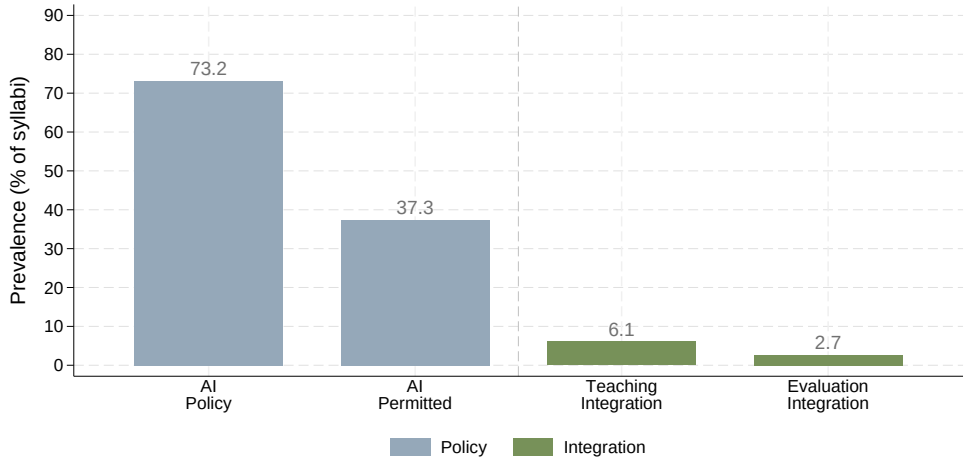
Figure 1 also shows that explicit integration of artificial intelligence into teaching is far less common than AI-focused policies. All in all, 6 percent of syllabi display concrete evidence of AI integration into teaching, whether through course content, class activities, or assigned readings. These results suggest that the main response to generative AI has taken the form of formal acknowledgment rather than direct integration into the classroom (McDonald et al., 2025).

AI integration into actual evaluation is even rarer in this context. By 2026, just 3 percent of courses include a graded task that requires students to engage with AI. Requiring students to engage with AI in graded work represents a distinct margin of integration. These

[9] In the full 2024–2026 sample, 62 percent of syllabi include an AI policy and close to 5 percent integrate it into their teaching. This reflects lower integration rates in the earlier semesters in the sample.

[10] Figure A3 shows the prevalence of these measures is broadly similar across the three classification models.

Figure 1: Fall 2026 Prevalence of the Four AI Measures



Note: Figure 1 shows the prevalence of the four AI measures in Fall 2026. The bars show the share of syllabi with each measure. In Fall 2026, 73.2% of syllabi include an AI policy, 37.3% explicitly permit AI use, 6.1% display teaching integration, and 2.7% display evaluation integration.

assignments may help students develop skills in using these tools, though they may also capture coursework that simply builds on AI-assisted work.

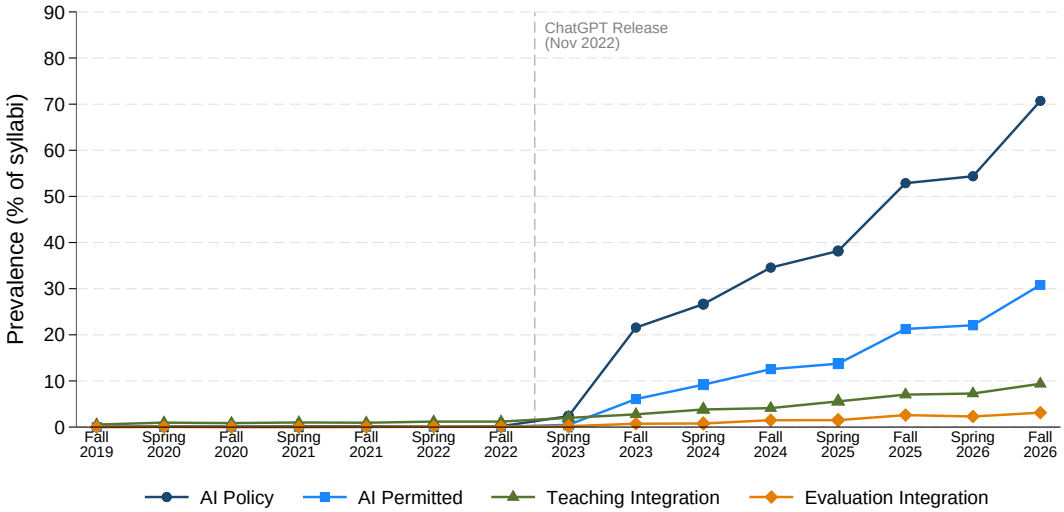
Qualitative Evidence. To further understand what these integration measures capture, I sort every syllabus classified as integrating AI into teaching or evaluation into a small set of descriptive categories. Rather than relying on another model judgment, the sorting applies fixed lists of keywords to the passage of syllabus text that the classifier quotes as evidence, which isolates the exact language describing how AI enters each course (Appendix C describes the approach).

On the teaching side, the most common category involves AI applied to the course's own field, accounting for 27 percent of the syllabi with teaching integration, while courses in which students work with AI tools and courses that cover AI methods each account for 18 percent. These categories also differ clearly in the language instructors use, as the terms most distinctive of AI Methods courses include machine learning and neural networks, while Using AI Tools courses are instead marked by references to ChatGPT and prompts, signaling that they capture distinct forms of integration. On the evaluation side, integration is dominated by graded tasks in which students produce their own work with AI assistance.[11]

AI Integration Over Time. Figure 2 shows how formal permission for AI use and classroom integration evolved at the four universities observed since Fall 2019, from Fall 2019 through Fall 2026. Reassuringly, I first find that no courses explicitly permit AI use prior to the rollout of ChatGPT. However, the share of syllabi with explicit permission increased

[11]This category accounts for 38 percent of the syllabi with evaluation integration, followed by tasks in which students build AI models and tasks in which students interact with or judge AI output.

Figure 2: AI Policy, AI Permitted, Teaching Integration, and Evaluation Integration Over Time at the Four Universities Observed Since Fall 2019



Note: Figure 2 shows the semester-by-semester evolution of AI Policy, AI Permitted, Teaching Integration, and Evaluation Integration from Fall 2019 through Fall 2026 at the four universities observed since Fall 2019 (Stephen F. Austin, Texas A&M, UT Austin, and UT Dallas). Each line gives the share of syllabi with the measure at the universities observed in that semester. In the semesters in which one university is not observed (Stephen F. Austin in Spring 2025, and Texas A&M in Spring 2023 and Spring 2024), the share covers the other universities. The vertical dashed line marks the launch of ChatGPT in November 2022.

rapidly over the following four years, so that close to 31 percent of courses at these universities now allow students to use generative AI. Classroom integration increased much more slowly, as teaching integration rose only to about 9 percent by Fall 2026.[12] Most of that integration still comes through the teaching margin, while evaluation remains comparatively rare throughout the period. Taken together, these patterns suggest that instructors adjusted formal permissions much faster than they changed what students were actually asked to do in class.

Table A4 presents the corresponding semester-by-semester changes for the four measures over the past five semesters. I find the same pattern in this broader sample, as the share of syllabi that explicitly permit AI use rises from 15 percent in Fall 2024 to 45 percent in Fall 2026. Classroom integration, in contrast, increases far more gradually over the same period, again driven by the teaching margin. Altogether, even in this shorter window, colleges appear to have moved faster in formalizing permission for AI than in incorporating it into classroom practice.

4.2 Heterogeneity Across Majors

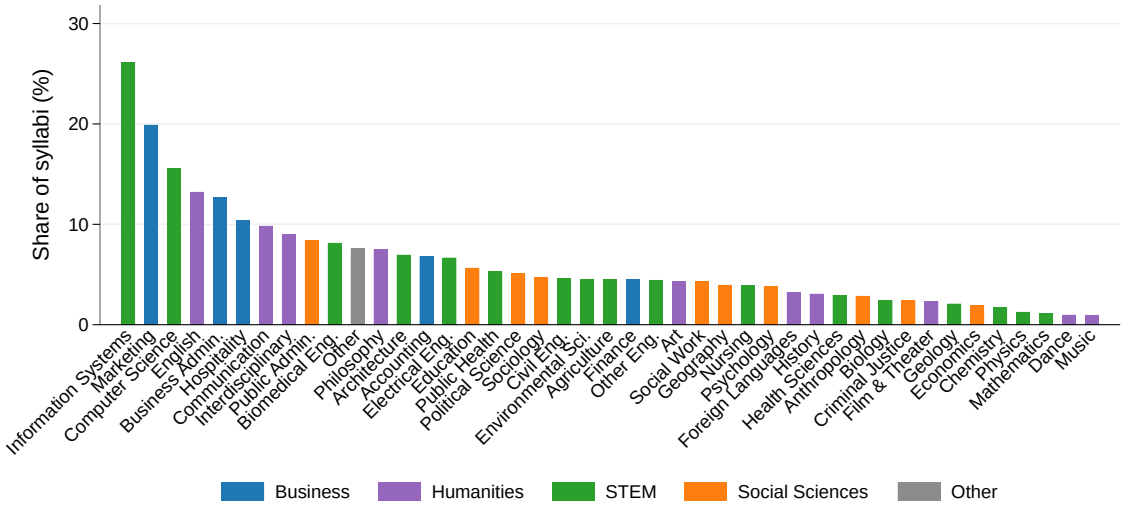
The results presented so far show that college professors have gradually started integrating AI into the classroom. In this context, as different fields prepare students for jobs

[12] A small share of courses already integrated AI into teaching before ChatGPT, largely because some syllabi discussed AI as a course topic, especially in computer science and business fields. This baseline reflects real course content rather than early use of generative AI tools.

with different exposure to generative AI (Eloundou et al., 2024), these patterns may also vary across majors (Contractor and Reyes, 2025; Handa et al., 2025).

Figure 3 presents evidence on the extent to which AI has been differentially integrated across college majors in Fall 2026. I find significant differences in adoption outcomes, as 20 percent of courses in Marketing have integrated AI into their teaching, compared to just one percent of courses in Music and Dance. Altogether, 9 percent of courses in Business majors have integrated AI into their teaching across the five semesters, which is roughly double the adoption rate for majors in the Humanities, STEM, and Social Sciences.[13] While this integration is more likely to occur through the teaching margin than through AI-related assignments (Figure A5), the two margins are strongly correlated across majors. Majors that are quicker to incorporate AI into course activities are also more likely to integrate it into graded assignments.

Figure 3: Teaching Integration Across Majors



Note: Figure 3 presents the share of syllabi with teaching integration in Fall 2026 across the harmonized majors, with bars colored by broad field (Business, Humanities, STEM, Social Sciences, Other). Bars are ordered from highest to lowest. Syllabi without a classified major are excluded.

On the other hand, courses in Social Sciences and in the Humanities are more likely to include an explicit AI policy in their syllabi (Table A5). This mirrors the differential prevalence of AI mentions across fields presented earlier. At the same time, this table shows that 34 percent of classes in Business majors explicitly allow students to use AI, which is higher than in the other majors. While instructors in Business classes are less likely to include an explicit AI policy, they are much more likely to permit students to use it when they do. Majors in Social Sciences and the Humanities have thus moved rapidly to adopt policies without making corresponding changes to integrate AI into the classroom.[14]

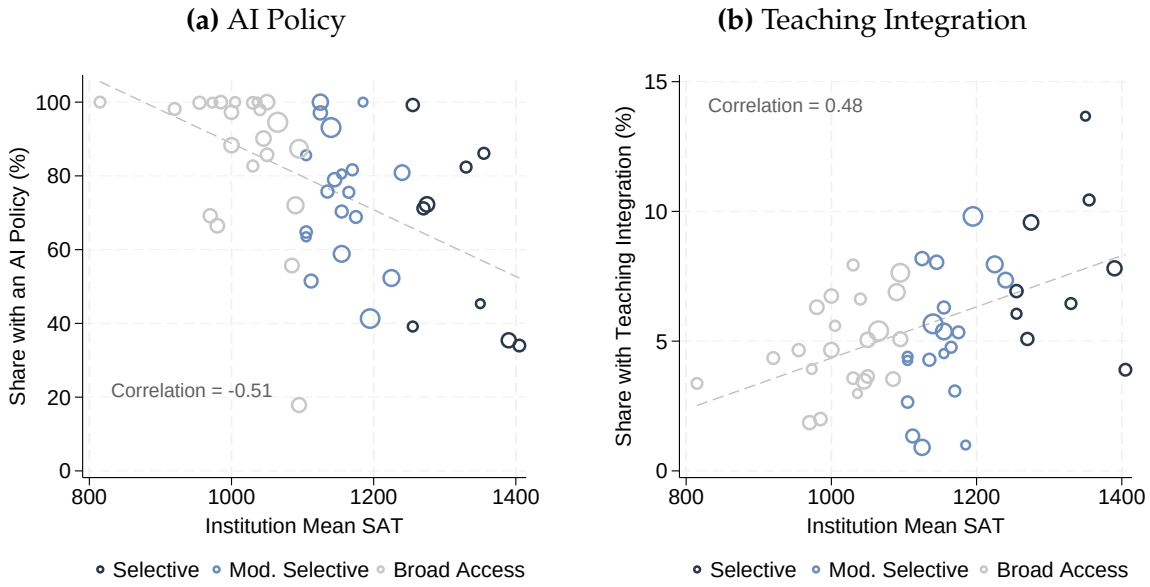
[13] While Business majors at the four universities observed since Fall 2019 were more likely to engage with AI content in the classroom prior to ChatGPT, the cross-field differences grow substantially after Fall 2022 (Figure A4).

[14] English classes stand out as an important exception, as this major is both more likely to have explicit AI policies and to have explicitly integrated it into the classroom.

4.3 Heterogeneity Across Universities

I also examine whether the prevalence of these measures varies significantly across the universities in the sample. The first panel of Figure 4 shows that some universities have incorporated AI policies in almost all of their courses, and that this is more likely to have happened at less selective universities. While every course at UT Permian Basin includes an AI policy in the Fall 2026 semester, this is the case in about a third of courses at the University of Florida (Table A6).

Figure 4: AI Integration and Mean SAT Score



Note: Figure 4 plots university-level averages of AI policy and teaching integration in Fall 2026 against mean SAT. Panel (a) plots AI policy prevalence against mean SAT. Panel (b) plots teaching integration prevalence against mean SAT. Marker size is proportional to the number of syllabi contributed by each institution and the dashed line is the corresponding syllabus-weighted linear fit. Marker color denotes the SAT selectivity tier. Mean SAT is the institution average of the 25th and 75th percentile scores for the 2019 entering class, or for the latest entering class reported where the 2019 value is missing.

On the other hand, the second panel of Figure 4 shows that more selective institutions are more likely to have incorporated generative AI into their teaching. About 8 percent of courses at selective institutions have already integrated AI, compared to around 5 percent at moderately selective and broad-access universities. In particular, while more than 10 percent of courses at UT-Austin and UT-Dallas have already integrated AI into their teaching, this share remains below 7 percent at West Texas A&M. Importantly, the selectivity gradient is mostly present on the teaching margin. This suggests that instructors' decisions around lecture materials and classroom activities may be shaping AI integration patterns across universities.

Altogether, these patterns are consistent with the growing evidence that larger, higher-paying, and better-resourced firms are more likely to adopt AI more deeply (McElheran et al., 2024; Yotzov et al., 2025). As such, a similar gradient is emerging in higher education as more selective institutions are far more likely to integrate it into the classroom.

4.4 University and Major Characteristics

The results presented so far show that the integration of AI into the classroom varies significantly across universities and college majors. To understand which universities have taken the lead on this front, I draw on the extensive set of university-level characteristics described in Section 2. In particular, I analyze how university resources and selectivity shape the prevalence of the main outcomes by estimating the following equation:

$$Y_n^{(k)} = X'_{u(n)}\beta_u^{(k)} + \phi_{m(n),l(n),t(n)} + \varepsilon_n^{(k)}, \qquad k \in \mathcal{K}, \tag{1}$$

where $Y_n^{(k)}$ is a binary variable measuring the prevalence of one of the AI-integration measures (k) in syllabus n. Meanwhile, $X_{u(n)}$ captures one of the main university-level characteristics, which include R1 status, research expenditures, average faculty salary, faculty per student, the tenured faculty share, spending and endowment per student, enrollment levels, mean SAT, the admission rate, the share of parents in the top income quintile, and the graduate enrollment share. Equation (1) includes major-by-course-level-by-term fixed effects ($\phi_{m(n),l(n),t(n)}$), allowing me to identify how AI adoption differs across universities within the same major, course level, and term. In practice, I estimate equation (1) separately for each outcome in $\mathcal{K}$, using one covariate at a time. I cluster standard errors at the university level.

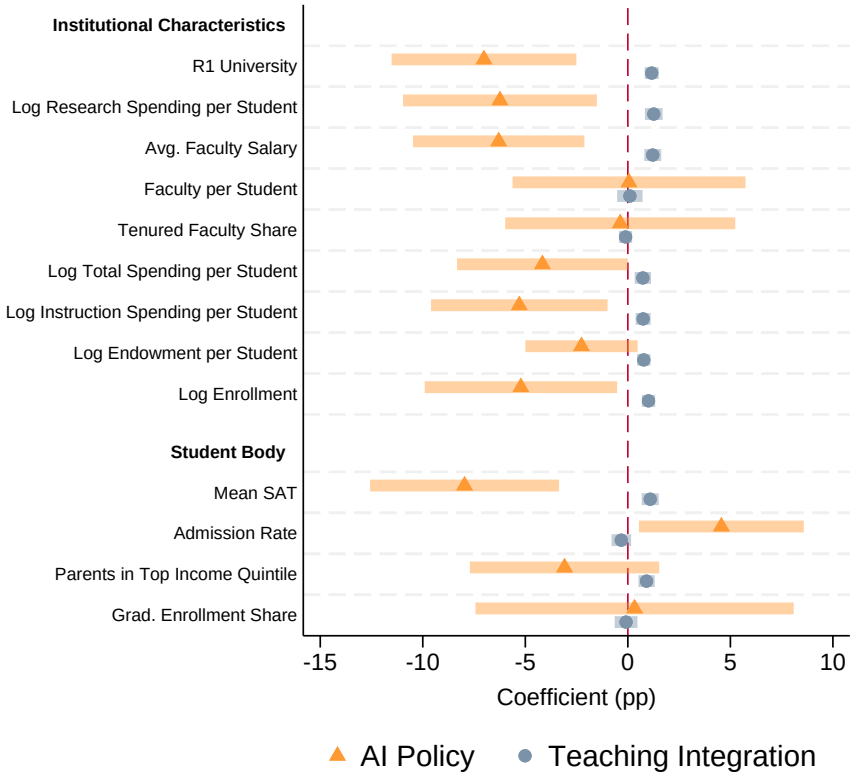
Moreover, given the sizable differences in adoption across majors documented above, I similarly examine which major characteristics are related to the main outcomes of interest. I draw on the main major-level covariates described in Section 2, which include the main measure of AI exposure and the concentration of the occupations that graduates enter, along with major-level mid-career earnings, unemployment rate, graduate-degree share, mean SAT, college GPA, and the share of female students in the field. To understand the drivers of AI integration at the major level, I estimate a modified version of equation (1), replacing the university-level covariate with a major characteristic ($X_{m(n)}$), while including university fixed effects along with course-level-by-term fixed effects. This specification compares the prevalence of AI integration for different majors within the same university and semester, thus allowing me to understand how these characteristics shape differential adoption across fields.

University-Level Results. Figure 5 presents the results from equation (1), showing how university-level characteristics shape the prevalence of AI policy language in syllabi. The results show that universities with lower average SAT scores are more likely to include AI policy language, confirming the evidence presented in Figure 4. Moreover, universities with lower average faculty salaries and lower research expenditures are similarly more likely to include policy language in their syllabi, as are universities without R1 status.

I also analyze how these characteristics shape the extent of teaching integration in Figure 5. First, R1 universities and those with higher research expenditures per student are more likely to integrate AI into their teaching. This pattern also emerges for institutions

with higher average faculty salaries, as a one-SD increase in salaries predicts a 1.2pp increase in teaching integration. These results suggest that universities with stronger research environments are more likely to integrate AI into their teaching.[15]

Figure 5: University Characteristics and AI Integration



Note: Figure 5 presents coefficient estimates from regressions of AI Policy and Teaching Integration on one university characteristic at a time, pooling the five semesters from Fall 2024 to Fall 2026. Each row is a separate regression. I standardize each characteristic to mean zero and unit standard deviation across universities, each counted once. The outcome is in percentage points. Bars represent 90% confidence intervals. Fixed effects are major by course level by term. Standard errors are clustered at the university level. I set missing characteristic values to the mean and include a missing-value indicator.

This evidence thus indicates that universities with stronger research environments are the most likely to have integrated AI into the classroom. This is consistent with Biasi and Ma (2022), who find that universities with more research-active faculty teach courses significantly closer to the knowledge frontier, even after controlling for other institutional characteristics. Moreover, the existing firm-level evidence similarly shows that more innovative companies are more likely to have integrated artificial intelligence (McElheran et al., 2024), as are firms that pay higher wages (Yotzov et al., 2025). These results suggest that a similar gradient may be emerging in higher education.

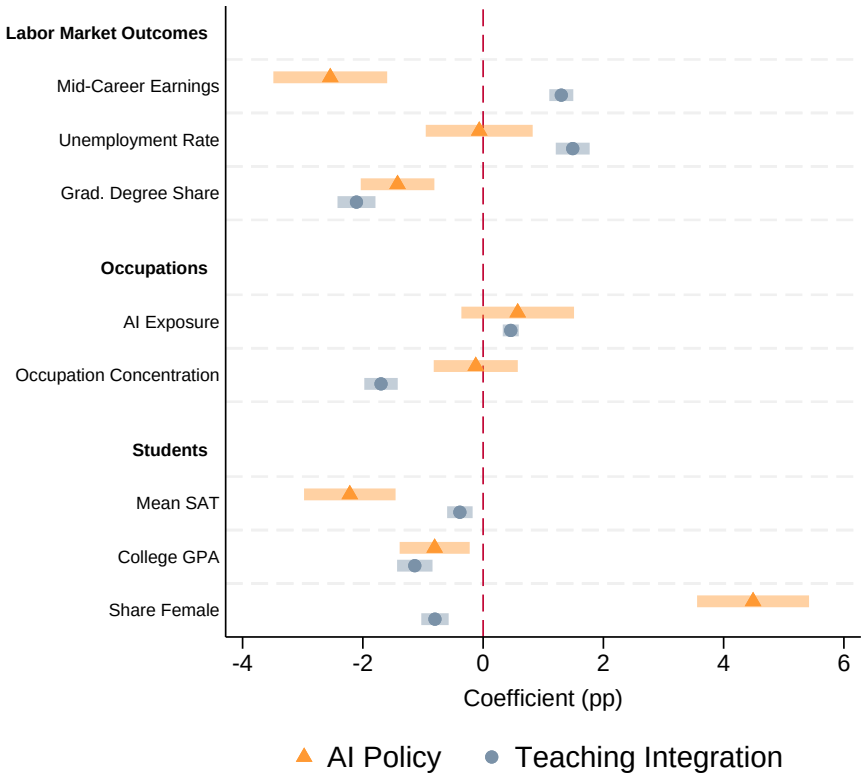
Major-Level Results. Figure 6 shows how major-level characteristics shape AI policy adoption conditional on university and course-level fixed effects. Majors with a larger share

[15] Institutions with higher mean SAT scores and a larger share of parents in the top income quintile are also more likely to integrate AI into their teaching. Figure A6 presents the corresponding results for AI permitted and evaluation integration.

of female students, lower graduate-degree shares, and lower mid-career earnings are more likely to have an explicit AI policy in place. By contrast, majors with greater exposure to AI are not significantly more likely to have put a policy in place.

Figure 6 also presents the corresponding results for teaching integration. First, majors with lower graduate-degree shares are more likely to integrate AI into their teaching, as are those with lower average SAT scores. I also find that majors with greater AI exposure are significantly more likely to have integrated these tools into their courses. AI exposure predicts AI-based teaching more strongly than evaluation (Figure A7). This fits in with the evidence presented in Section 4.2, where the fields with the highest rates of classroom integration are also those where students and workers are most likely to use generative AI.

Figure 6: Major Characteristics and AI Integration



Note: Figure 6 presents coefficient estimates from regressions of AI Policy and Teaching Integration on one major characteristic at a time, pooling the five semesters from Fall 2024 to Fall 2026. Each row is a separate regression. I standardize each characteristic to mean zero and unit standard deviation across the harmonized majors, each counted once. The outcome is in percentage points. Bars represent 90% confidence intervals. Fixed effects are university and course level by term. Standard errors are clustered at the university level. I set missing characteristic values to the mean and include a missing-value indicator.

4.5 Course Characteristics

In Table A7, I further examine the extent to which AI has been differentially integrated across different course levels. Explicit policy language is more common in undergraduate

courses than in graduate courses.[16] By contrast, graduate classes are most likely to have integrated AI into teaching. Altogether, 8 percent of graduate courses had incorporated AI into their teaching by the Fall of 2026. In this sense, graduate courses resemble the most selective universities in the sample, with less AI policy language but more actual classroom integration.

To understand whether these patterns hold when accounting for differences across institutions, I estimate a modified version of equation (1). This specification regresses the main outcomes on observable course characteristics while including university-by-major-by-term and instructor fixed effects. I thus compare different courses taught by the same instructor while accounting for the broader teaching context in that term. As such, I analyze how enrollment levels, credit hours, and course levels shape AI integration outcomes.

Figure 7 shows that AI policy language is concentrated in introductory courses. Conditional on instructor and university-by-major-by-term fixed effects, freshman courses are about 4 percentage points more likely than graduate courses to include an AI policy, and this difference narrows at each successive undergraduate level. Larger courses and courses with teaching assistants are also more likely to include AI policy language, fitting in with the broader pattern that these high-enrollment settings have been the primary site of the institutional response to AI.

On the other hand, graduate-level courses are more likely than advanced undergraduate classes to have actually integrated AI into the classroom. This difference between the policy and integration results is especially clear on the teaching side, where advanced undergraduate courses are significantly less likely than graduate courses to dedicate class content to AI. In this context, graduate courses may have more natural scope for incorporating AI into teaching, precisely because their content engages with frontier knowledge (Biasi and Ma, 2026).[17]

5 Drivers of AI Integration

5.1 Variance Decomposition

To understand the extent to which instructors, courses, and the broader academic environment contribute to shaping AI integration, I estimate the following equation:

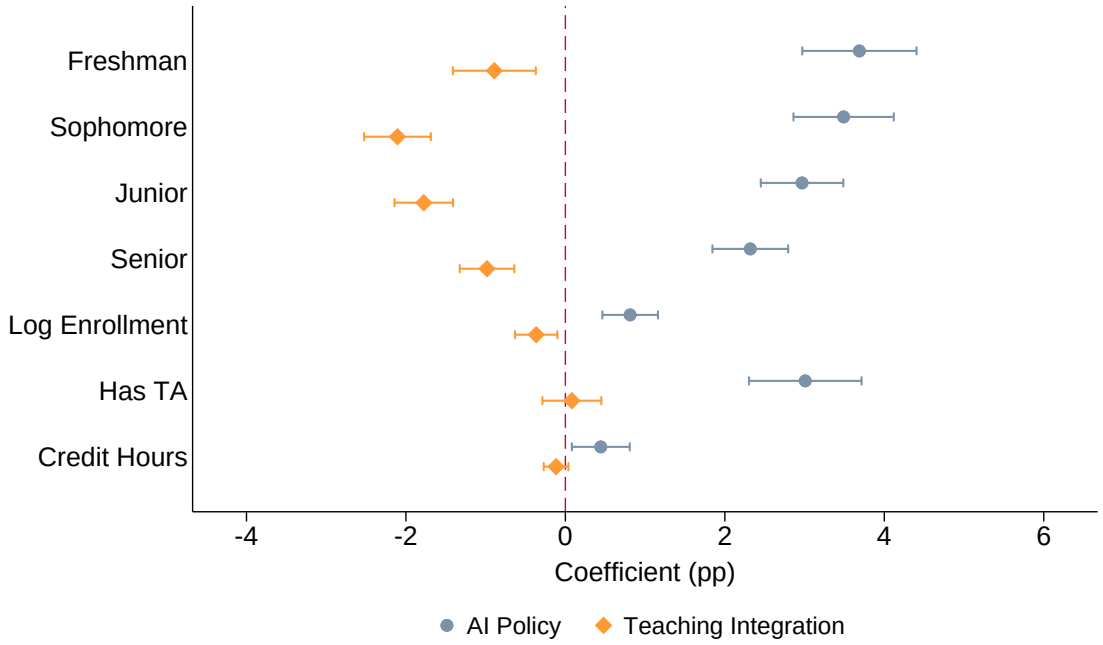
$$Y_n^{(k)} = \alpha_{i(n)} + \lambda_{c(n)} + \phi_{u(n),m(n),l(n),t(n)} + \varepsilon_n^{(k)}, \qquad k \in \mathcal{K}, \tag{2}$$

where $\alpha_{i(n)}$ captures instructor fixed effects, whereas $\lambda_{c(n)}$ denotes local course fixed effects, representing university-specific courses. These course effects absorb stable features of a given course, such as its recurring content, structure, or assessment design, that may shape the scope for AI integration. Meanwhile, $\phi_{u(n),m(n),l(n),t(n)}$ denotes university-by-major-by-

[16] More than 60 percent of both basic and advanced undergraduate syllabi contain an AI policy, compared with 58 percent of graduate syllabi.

[17] At the universities with earlier terms, courses that were already offered before Fall 2024 are more likely to include an AI policy, but less likely to have integrated AI into teaching and evaluation (Table A8).

Figure 7: Course Characteristics and AI Integration



Note: Figure 7 shows the multivariate course-characteristic coefficients for AI Policy and Teaching Integration. The specification replaces local course fixed effects with observable course characteristics, includes instructor fixed effects and university-by-major-by-term fixed effects, and omits Graduate courses as the reference category. Missing indicators are included for Log Enrollment and Credit Hours. Credit hours and teaching-assistant indicators use the available course records. Enrollment is unavailable for Fall 2026 because schedule enrollment did not reproduce the earlier measure. Bars represent 90% confidence intervals. Standard errors are two-way clustered by instructor and local course.

course-level-by-term fixed effects, which absorb the broader institutional and disciplinary context in which the syllabus is written. In practice, the decomposition is informed by comparisons across syllabi that share part of the same academic environment while differing along other dimensions. For instance, when the same local course is taught by different instructors, the remaining variation helps recover the instructor margin. Similarly, when instructors appear across different courses or terms within the same broader environment, the specification reveals how much of the variation is tied to the course itself versus the surrounding institutional and disciplinary context.[18]

To further understand whether AI integration is driven by individual instructors, by the courses they teach, or by the universities and majors in which they operate, I implement a Shapley–Owen decomposition (Audoly et al., 2025). In particular, this decomposition computes the average marginal contribution of each of these components to adjusted R^2 across all possible orderings of inclusion.[19] The Shapley–Owen decomposition thus recovers the partial contributions of instructors, courses, universities, and majors to AI integration in an

[18] I also estimate a more stringent version that replaces $\lambda_{c(n)}$ with local-course-by-year fixed effects (λ_{cy}), tightening the comparison to instructors teaching the same course in the same academic year.

[19] Formally, for a given block j, the Shapley contribution equals $\sum_{T \subseteq V \setminus \{j\}} \frac{|T|!(K-|T|-1)!}{K!} \left[R^2(T \cup \{j\}) - R^2(T)\right]$, where $R^2(S)$ denotes the adjusted R^2 from the specification that includes the subset of blocks S, V is the full set of blocks, and $K = |V|$.

additive and order-invariant approach.

Table 3 presents the results from the variance decomposition of equation (2) for the main measures of AI integration. These results show that AI policy is shaped largely by instructors. That is, instructors explain about 31 percent of the variance in whether a syllabus includes AI policy language, while local course fixed effects account for an additional 21 percent of the variation. I find a similar pattern for whether AI is explicitly permitted. There too, instructors explain the largest share of the variance, local courses explain a smaller but still important share, and the broader environment accounts for relatively little.

Table 3: Variance Decomposition of AI Integration

	AI Policy	AI Permitted	Teaching Integration	Evaluation Integration
	(1)	(2)	(3)	(4)
Instructor	30.7	33.4	24.5	23.6
Local Course	20.8	19.6	31.8	25.4
University	5.1	2.5	0.3	0.1
Major	1.1	0.7	1.4	0.4
Term	6.9	5.0	0.2	0.2
Unexplained	35.4	38.9	41.8	50.3
Adjusted R^2	64.6	61.1	58.2	49.7
Observations	250,711	250,711	250,677	250,821
Universities	50	50	50	50

Note: Table 3 presents the Shapley–Owen variance decomposition for AI Policy, AI Permitted, Teaching Integration, and Evaluation Integration. The five components are instructor, local course, university, major, and term fixed effects. I calculate each share using the same outcome-specific rows in every subset regression. Each entry gives the share of total outcome variance in percentage points. Adjusted R^2 is the total explained share. The sample pools the syllabi from Fall 2024 to Fall 2026 with an identified instructor, course and major, and excludes syllabi that are the only observation of their instructor or course.

I find similar patterns on the integration margins. Instructors explain around one-fourth of the variation across the integration measures, with local course fixed effects accounting for a comparable share.[20] By contrast, the university-by-major-by-level-by-term component explains a very small share of the variation in AI integration.[21] This indicates that professors and the courses they teach drive most of the systematic variation in whether AI is integrated into higher education.[22]

Table 3 presents the estimates from the full Shapley–Owen decomposition. Altogether, instructors and local courses remain the primary drivers of AI policy and actual integration. At the same time, universities account for 5 percent of the variation in AI policy, but a negligible amount of the variation in whether it is actually integrated into teaching or evaluation. Universities thus shape whether AI is formally regulated in the syllabus but do not contribute to whether it is actually incorporated into teaching. This pattern is consistent

[20] When including course-year fixed effects, instructors explain the largest share of the variance in AI policy and AI permitted, and a slightly smaller share than courses for teaching and evaluation integration.

[21] Figure A8 shows that the relative magnitudes of these components are similar across the five broad major groups and university selectivity tiers.

[22] Table A9 shows that part of the course component reflects broader course characteristics that extend across institutions, rather than university-specific implementation.

with Biasi and Ma (2022), who similarly find that instructors and courses are the primary drivers of the integration of frontier research into teaching.

5.2 Instructor Changes

Empirical Approach. Since the decomposition does not reveal whether AI integration follows the instructor when a course changes hands, I compare consecutive offerings of the same course at a given university, following Biasi and Ma (2026).[23] In some cases, a new instructor takes over the course in the second offering, whereas in others the same instructor teaches both offerings. I then compare the change in AI integration between the two offerings in courses that change instructors with the change in courses whose instructor stays. This comparison accounts for the growth in AI integration that is common to all courses at the same university. If AI integration depends on who teaches the course, it should rise when an instructor who integrates AI takes over, and fall when such an instructor leaves.

To assess whether instructors bring their prior AI integration with them, I classify each instructor by whether they had integrated AI into their teaching in other courses. In particular, I define an integrator as an instructor who integrated AI into teaching in at least one syllabus of another course taught before the change. Since this definition relies on instructors' other courses, I focus on changes in which both instructors had taught another course before the change. An integrator enters a course when a non-integrator hands it over to an integrator, and an integrator leaves in the opposite case. Lastly, the remaining changes take place between two instructors of the same type. Altogether, I estimate the following difference-in-differences regressions:

$$Y_{ct} = \lambda_c + \phi_{u(c),\tau(c),t} + \beta\, D_c \times Post_t + \varepsilon_{ct} \tag{3}$$

$$Y_{ct} = \lambda_c + \phi_{u(c),\tau(c),t} + (\beta_E E_c + \beta_L L_c + \beta_S S_c) \times Post_t + \varepsilon_{ct} \tag{4}$$

where Y_{ct} is the share of course c's syllabi that integrate AI into teaching in offering t, in percentage points, and $Post_t$ indicates the second of the two consecutive offerings. Moreover, D_c indicates whether a different instructor teaches the second offering, so that β captures the change in AI integration in courses that change instructors relative to the change in courses whose instructor stays. In equation (4), E_c, L_c and S_c indicate whether an integrator enters, an integrator leaves, or the change takes place between two instructors of the same type. The course fixed effects λ_c absorb the fixed features of the course, such as its subject and level. In addition, $\phi_{u(c),\tau(c),t}$ denotes university-by-term fixed effects, where $\tau(c)$ captures the two terms in which the course is offered. These fixed effects allow me to compare each instructor change with courses at the same university whose instructor stays over the same two terms. In the more stringent specifications, I instead include university-by-major-by-term fixed effects, which restrict this comparison to courses in the same major. I cluster standard errors at the university level.

[23] I focus on offerings taught by a single instructor. I also exclude section rotations, in which the incoming instructor had already taught a section of the course.

The integrator types in equation (4) rely on instructors' prior AI integration in their other courses. However, instructors who had integrated AI before the change may have taught courses that are more prone to AI integration, even within the same major. As a result, the integrator label may partly capture the courses that instructors had taught, rather than the instructors themselves. To address this concern, I draw on the teacher value-added literature and predict each instructor's AI integration relative to colleagues who teach similar courses (Rockoff, 2004; Kane and Staiger, 2008; Chetty et al., 2014). In particular, I compare the syllabi that instructor i wrote for each other course j in earlier terms with those of these colleagues, and combine these differences into a prediction of the instructor's AI integration:

$$\hat{\mu}_i = \kappa_i \sum_j \omega_{ij} \left(\bar{Y}_{ij} - \bar{Y}_{ij}^{-c} \right) \tag{5}$$

where $\bar{Y}_{ij}$ is the share of instructor i's syllabi in course j that integrate AI into teaching, and $\bar{Y}_{ij}^{-c}$ is the average among syllabi at the same university, major, course level and term, excluding the course that changes instructors. The weights ω_{ij} increase with the number of syllabi in each course, whereas the factor κ_i shrinks the prediction toward zero when it rests on few syllabi (Walters, 2024). The factor also calibrates the prediction, so that a one percentage point increase in $\hat{\mu}_i$ predicts a one percentage point increase in AI integration in the instructor's later courses.[24]

I then estimate how much of the difference between the incoming and outgoing instructors' predicted AI integration carries over to the course:

$$Y_{ct} = \lambda_c + \phi_{u(c),\tau(c),t} + \left[\theta D_c + \delta D_c \left(\hat{\mu}_c^{in} - \hat{\mu}_c^{out} \right) + \gamma (1 - D_c) \hat{\mu}_c^{own} \right] \times Post_t + \varepsilon_{ct} \tag{6}$$

where $\hat{\mu}_c^{in}$ and $\hat{\mu}_c^{out}$ denote the predicted AI integration of the incoming and outgoing instructors, and $\hat{\mu}_c^{own}$ that of an instructor who keeps the course. The coefficient δ thus captures the share of the difference between the two instructors that carries over to the course, which would equal one if instructors brought all of their AI integration with them to the new course. The last term allows courses taught by instructors with higher predicted AI integration to evolve differently, even when their instructor stays.

Main Results. Table 4 presents the difference-in-differences estimates from equations (3) and (4). First, I find that an instructor change has little effect on AI integration on average. Relative to courses whose instructor stays, AI integration rises by only 0.35 percentage points when a new instructor takes over, and this estimate remains small across specifications. This differs from the evidence on frontier knowledge, as Biasi and Ma (2026) show that courses move closer to the research frontier when a new instructor takes over. By contrast, a new instructor does not necessarily bring AI into the classroom.

However, this average may mask differences across the instructors who take over these courses. I therefore examine whether new instructors differentially integrate AI based on

[24] In courses held out from this calibration, a one percentage point increase in the prediction corresponds to a one percentage point increase in AI integration.

Table 4: Instructor Changes and AI Integration

	(1)	(2)	(3)	(4)
Panel A. All Instructor Changes				
Instructor Change	0.35*	0.41*	0.37*	-0.76
	(0.20)	(0.21)	(0.22)	(0.65)
Panel B. By Instructor Type				
Integrator Entered	9.51***	9.22***	9.15***	8.17***
	(1.35)	(1.32)	(1.32)	(2.87)
Integrator Left	-6.33***	-6.41***	-6.43***	-11.41***
	(1.67)	(1.63)	(1.64)	(2.89)
Same	0.13	0.22	0.19	-0.33
	(0.19)	(0.20)	(0.21)	(0.61)
Stayer Baseline	1.03	1.03	1.03	1.08
University × Semester Pair FE	Yes	No	No	No
University × Field × Semester Pair FE	No	Yes	Yes	Yes
Instructor Rank	No	No	Yes	No
Seen Two Offerings Before	No	No	No	Yes
Changes	12,606	12,606	12,606	1,848
Stayer Pairs	42,438	42,438	42,438	15,089
Universities	39	39	39	33

Note: Table 4 presents OLS estimates of the change in the share of a course's syllabi that integrate AI into teaching, in percentage points, from the last offering before an instructor change to the first offering after it. Each regression compares courses that change instructor with courses at the same university over the same pair of semesters whose instructor stays. Column 2 compares courses within the same field, column 3 adds the academic rank of both instructors, and column 4 keeps the changes whose course is also observed two offerings before the change. The stayer baseline is the mean change among courses whose instructor stays. Standard errors clustered by university in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

their experience in their previous courses. The second panel of Table 4 shows that when an integrator takes over a course from a non-integrator, the likelihood that the course integrates AI into teaching increases by 9.5 percentage points relative to courses whose instructor stays.[25] That is, courses that gain an integrator more than double their AI integration. On the other hand, AI integration falls by 6.3 percentage points when an integrator leaves, whereas changes between two instructors of the same type leave AI integration unchanged. These estimates remain similar when comparing courses within the same major and when accounting for the academic rank of both instructors. As such, the small average effect of an instructor change combines a rise in AI integration when an integrator arrives with a decline when an integrator leaves.

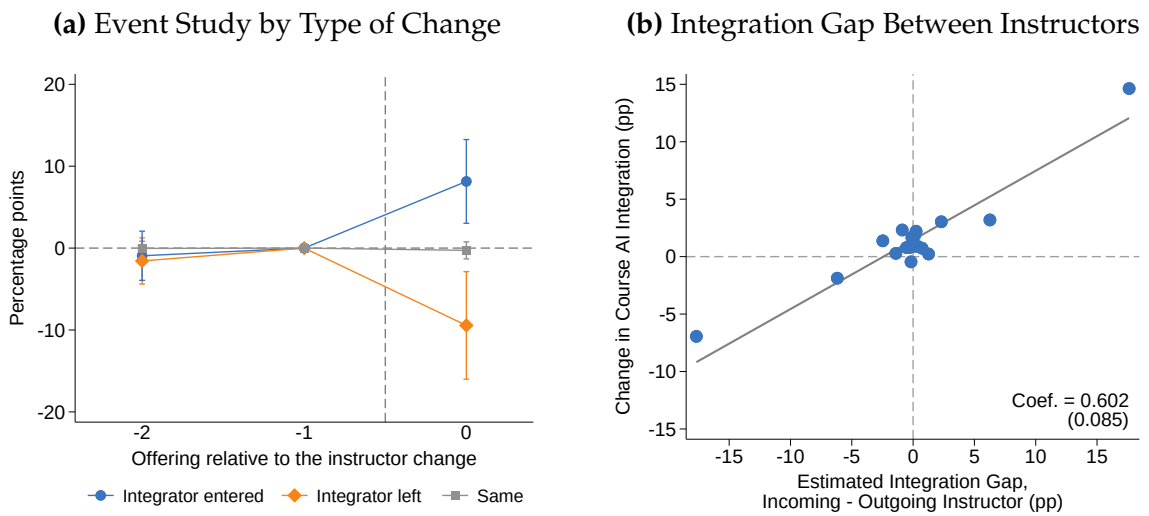
Importantly, the last column of Table 4 shows that these results are robust to restricting the sample to courses that the outgoing instructor had taught in at least two offerings before the change. To ensure the results are not driven by pre-existing trends in AI integration, I estimate an event-study version of equation (4) on courses observed in two offerings before the change, and present it in the first panel of Figure 8. These estimates show no evidence of pre-trends in AI integration for any of the three types of instructor change. AI integration then rises sharply in the first offering taught by an incoming integrator, and falls when an integrator leaves.[26] The change in AI integration thus coincides with the arrival of the new

[25] In raw terms, AI integration in these courses rises from 6 to 17 percent, compared with an increase from 4 to 5 percent in courses whose instructor stays.

[26] Despite the smaller sample size, Figure A9 similarly shows no rise in AI integration ahead of the change

instructor.

Figure 8: Instructor Changes and AI Integration



Note: Figure 8 presents the change in AI integration around instructor changes. Panel (a) presents estimates from an event-study version of equation (4) on courses observed in two offerings before the change, with the last offering before the change as the reference, relative to courses at the same university whose instructor stays. Bars represent 95% confidence intervals based on standard errors clustered by university. Panel (b) groups instructor changes into 20 equal-sized bins by the difference between the incoming and outgoing instructors' predicted AI integration, and plots the average change in AI integration in each bin, both net of the other terms in equation (6). The line and its slope, with the standard error clustered by university in parentheses, come from estimating equation (6) on the underlying data.

Lastly, Table A10 shows that instructors who had previously integrated AI into their teaching are also more likely to allow students to use AI and to integrate AI into graded work when they take over a new course. In particular, courses that gain an integrator become 7.6 percentage points more likely to explicitly permit AI use and 2.8 percentage points more likely to integrate AI into graded work, which only 1.5 percent of courses did before the change. These results suggest that instructors who integrate AI into their teaching also shape how students engage with AI in the courses they take over.

The results so far rely on the integrator label, which may partly reflect the courses that instructors had taught before the change. The second panel of Figure 8 therefore presents the estimates from equation (6), which relies on the instructors' predicted AI integration instead. The figure shows that courses integrate AI more when the incoming instructor's predicted integration exceeds that of the outgoing instructor, and less in the opposite case. In fact, the estimated slope indicates that a course takes on 60 percent of the difference between the two instructors' predicted AI integration.[27] This provides further evidence that instructors bring their AI integration with them, even after accounting for the courses they had taught before the change. These results fit in with Chetty et al. (2014) and Bacher-Hicks et al. (2014), who show that changes in the teaching staff of a school shift student test scores in line with teachers' value-added.

Altogether, the evidence from instructor changes shows that whether a course integrates

for courses observed in three offerings before the instructor change.

[27] The slope is similar when the prediction excludes all courses that change instructors, and when the outgoing instructor leaves the university rather than teaching other courses there.

AI depends on who teaches it. In particular, instructors who had integrated AI in their other courses bring it with them when they take over a course, and AI integration declines when they leave. Since instructors play a central role in whether AI enters the classroom, I next examine which instructors lead this adoption.

6 Which Instructors Integrate AI?

6.1 Instructor Characteristics

The results presented in Section 5 show that instructors play a key role in shaping how AI is incorporated into the classroom. In particular, AI integration rises in courses that an integrator takes over and falls in those that an integrator leaves. However, these results do not reveal which instructors are more likely to integrate AI in the first place.

Figure 9 presents descriptive evidence on how teaching integration and AI policy vary across instructor characteristics. First, the prevalence of explicit policies mirrors the results on AI mentions presented in Section 2. That is, women are more likely to include explicit policies in their syllabi than men, as 67 percent of syllabi written by women include an AI policy, compared with 58 percent of those written by men. Full professors are also the least likely to include a policy in their syllabi.

At the same time, there are important differences in which instructors actually integrate AI into the classroom. On this margin, the gender difference largely disappears. Moreover, actual integration is more common among assistant professors, as 5.7 percent of their syllabi integrate AI into teaching, compared with 4.8 percent of those written by full professors. Similarly, professors who have published papers are more likely to have AI in their teaching than their peers with no publications, even though they are less likely to include an AI policy.[28] In that sense, the instructors who are quickest to formalize AI in the syllabus are not the same as those most likely to incorporate it into their courses.

6.2 Main Analysis

Empirical Framework. To assess the extent to which instructor characteristics shape AI outcomes conditional on the courses they teach and the institutional context in which they operate, I estimate the following regression:

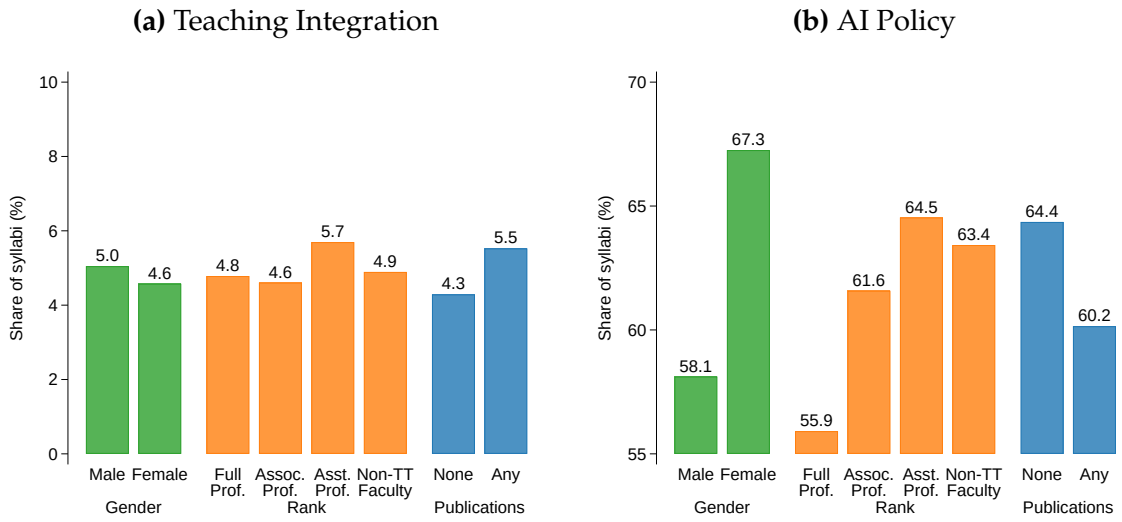
$$Y_n^{(k)} = X'_{i(n)}\beta^{(k)} + \lambda_{c(n)} + \phi_{u(n),m(n),l(n),t(n)} + \varepsilon_n^{(k)}, \qquad k \in \mathcal{K}, \tag{7}$$

where $X_{i(n)}$ includes instructor gender, academic rank relative to full professors, an indicator for any publications, and the years since the instructor's last publication.[29] The fixed effects $\lambda_{c(n)}$ and $\phi_{u(n),m(n),l(n),t(n)}$ absorb local-course and university-by-major-by-course-level-by-term variation, respectively, so that the estimates compare instructors teaching the

[28] Figure A10 shows that whether a syllabus explicitly permits AI follows similar patterns to AI policy, while AI in graded work varies little across these groups.

[29] The years since the last publication are measured in standard deviations among instructors with publications and set to zero for instructors without publications. I also include indicators for missing characteristics.

Figure 9: Teaching Integration and AI Policy by Instructor Characteristics



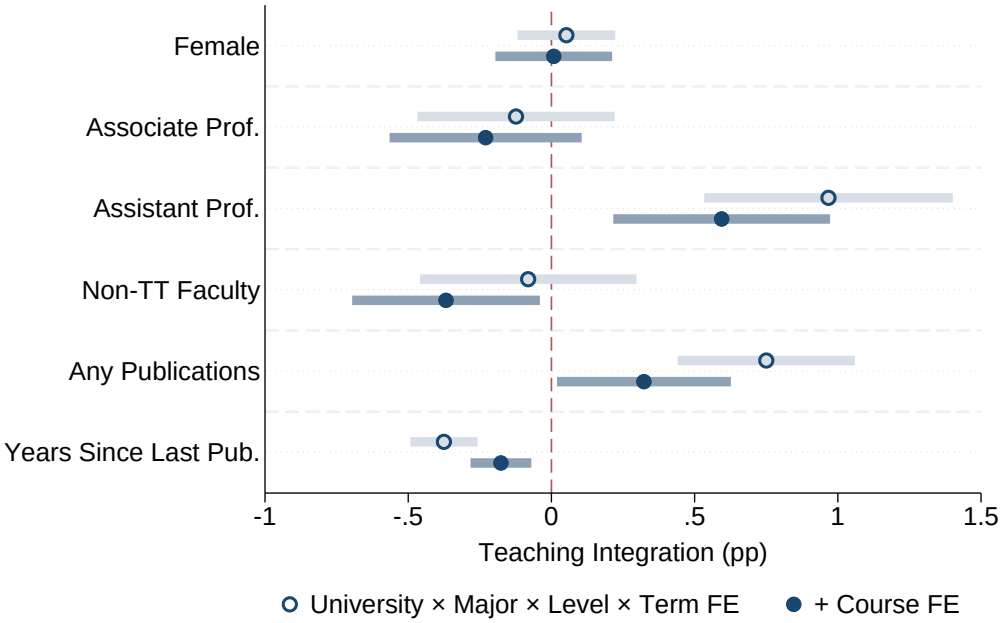
Note: Figure 9 presents the share of syllabi that integrate AI into teaching (panel a) and that include an explicit AI policy (panel b), in percent, by the gender, academic rank and publication record of the instructor. Each group includes the syllabi whose instructor has that characteristic recorded. The sample covers the 50 universities from Fall 2024 through Fall 2026.

same course in the same context. I also estimate a version of equation (7) without the course fixed effects, which compares instructors in the same university, major, course level and term who may teach different courses. The difference between the two sets of estimates shows the extent to which each characteristic reflects the courses that instructors teach. I cluster standard errors at the university level.

Empirical Results. Figure 10 presents the estimates for teaching integration. I find that assistant professors are more likely than full professors to integrate AI into their teaching, even conditional on the courses they teach. In particular, assistant professors are 0.59 percentage points more likely to integrate AI than full professors teaching the same course, a difference amounting to 12 percent of the sample mean. Instructors with published papers are also more likely to have incorporated AI into their courses, though this difference declines with the time since their most recent publication. These results thus suggest that young tenure-track faculty with active research agendas are leading the integration of AI in the classroom, fitting in with evidence that younger and more educated workers tend to lead the adoption of artificial intelligence in the workplace (McElheran et al., 2024; Bick et al., 2025).

Importantly, part of these differences reflects the courses that instructors teach. When I compare instructors in the same university, major, course level and term, assistant professors are 0.97 percentage points more likely than full professors to integrate AI, and this gap narrows once I compare instructors teaching the same course. The course fixed effects absorb an even larger share of the gap between instructors with and without publications, as instructors with publications are more likely to teach courses in which AI integration is common. Nevertheless, both differences remain among instructors teaching the same

Figure 10: Instructor Characteristics and Teaching Integration



Note: Figure 10 presents coefficient estimates from regressions of teaching integration, measured in percentage points, on instructor characteristics entered jointly. Hollow circles show estimates from the specification with university-by-major-by-level-by-term fixed effects. Solid circles add course fixed effects. Both specifications are estimated on the same sample. Full professors are the omitted rank. Years Since Last Pub. is the number of years since the instructor's most recent publication in standard deviations among instructors with publications and equals zero for instructors without publications. Standard errors are clustered at the university level. The sample covers the 50 universities from Fall 2024 through Fall 2026. Bars represent 90% confidence intervals.

course. I also find that these estimates are unchanged when I control for whether the instructor is teaching the course for the first time, which requires designing new materials rather than updating existing ones.[30]

On the other hand, non-tenure-track faculty are less likely than full professors to integrate AI into their teaching, a difference that only emerges among instructors teaching the same course. Since non-tenure-track faculty are hired primarily to teach (Figlio et al., 2015), the time available for teaching does not by itself explain which instructors integrate AI. Lastly, women and men are equally likely to integrate AI into the courses they teach.

Table 5 further examines which aspects of an instructor's career and research shape these patterns. First, column 2 adds the number of years since the instructor's first publication, which captures how far along instructors are in their research careers. Instructors who began publishing more recently are more likely to integrate AI into their teaching, and the assistant professor gap falls to 0.45 percentage points once I account for this measure. At the same time, the years since the last publication remain an important predictor of AI integration. As such, both an early career stage and recent research activity shape which instructors integrate AI into their teaching.

Moreover, columns 3 and 4 show that professors who publish more or are more widely

[30] I observe whether an instructor had taught the course before at the 39 universities with syllabi from more than one term.

Table 5: Instructor Characteristics and Teaching Integration: Career Stage and Research Output

	Main	Career Stage	Publications	Citations
	(1)	(2)	(3)	(4)
Assistant Prof.	0.59**	0.45**	0.62**	0.65**
	(0.23)	(0.22)	(0.24)	(0.25)
Associate Prof.	-0.23	-0.30	-0.21	-0.20
	(0.20)	(0.20)	(0.20)	(0.20)
Non-TT Faculty	-0.37*	-0.50***	-0.34*	-0.32
	(0.20)	(0.18)	(0.19)	(0.19)
Any Publications	0.32*	0.21	0.40*	0.36*
	(0.18)	(0.17)	(0.21)	(0.20)
Years Since Last Pub.	-0.18***	-0.17**	-0.18**	-0.16**
	(0.06)	(0.07)	(0.07)	(0.07)
Female	0.01	-0.01	0.01	0.02
	(0.12)	(0.12)	(0.12)	(0.12)
Years Since First Pub.		-0.25*		
		(0.12)		
Top Publication Quartile			0.02	
			(0.30)	
Top Citation Quartile				0.17
				(0.27)
Joint Test of Quartiles (p-value)			0.54	0.54
Outcome Mean	4.89	4.89	4.89	4.89
Course FE	Yes	Yes	Yes	Yes
University × Major × Level × Term FE	Yes	Yes	Yes	Yes
Observations	276,284	276,284	276,284	276,284
Universities	50	50	50	50

Note: Table 5 presents OLS regression estimates of teaching integration, in percentage points, on instructor characteristics. Full professors are the omitted rank. Years Since Last Pub. and Years Since First Pub. are measured in standard deviations among instructors with publications and equal zero for instructors without publications. Column 2 adds the years since the instructor's first publication. Columns 3 and 4 add indicators for the quartile of the instructor's publication count and citation count within their field, among instructors with publications, with the bottom quartile omitted. The table shows the top quartile, and the joint test covers the second, third and top quartiles. All columns absorb course and university-by-major-by-level-by-term fixed effects. The sample covers the 50 universities from Fall 2024 through Fall 2026. Standard errors clustered by university in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

cited are no more likely to integrate AI into their teaching than other instructors with publications. In both cases, I cannot reject that AI integration is the same across the four quartiles of publications or citations. However, Table A11 shows that AI integration rises with publications and citations when I compare instructors across courses in the same university, major, course level and term. In other words, instructors who publish more are more likely to teach the courses in which AI is integrated, but they are no more likely to integrate it than other instructors with publications who teach the same course. This differs from the evidence on frontier knowledge, as Biasi and Ma (2022) show that instructors who publish more and are more cited teach content closer to the research frontier in their own fields.

Instead, Table A12 shows that the gap between instructors with and without publications is concentrated among those who published in the year of the course or the year before. By contrast, instructors whose last publication came out earlier are no more likely to integrate AI than those without publications. Therefore, the recency of an instructor's re-

search plays a larger role than its volume in whether they bring AI into the classroom. This fits in with prior work showing that individuals learn to use new technologies through their own experience and that of their peers (Foster and Rosenzweig, 1995).

To examine whether these patterns reflect instructors' attention to AI more broadly, Figure A11 presents the corresponding estimates for AI policy and AI permitted. On the policy margin, women are 3.7 percentage points more likely than men to include an explicit AI policy in their syllabi, and non-tenure-track faculty are similarly more likely to do so than full professors. Assistant professors are also more likely than full professors to have a policy in place, whereas instructors with publications are no more likely to do so.[31] Since women and non-tenure-track faculty are no more likely to integrate AI into their teaching, attention to AI in the syllabus does not by itself explain which instructors bring it into the classroom.

On the evaluation margin, instructors with publications are also more likely to integrate AI into graded work, a difference amounting to 16 percent of the sample mean (Figure A12). At the same time, assistant professors are no more likely than full professors to have done so. The career-stage differences arise mainly through the teaching margin rather than through the inclusion of AI in graded work.

The main results remain consistent across various robustness checks. First, Table A13 shows that the estimates are similar when I replace the course fixed effects with course-by-academic-year fixed effects, which compare instructors teaching the same course in the same academic year. Moreover, the assistant professor gap is smaller but remains when I control directly for the text-based characteristics described in Section 2, which account for differences in how instructors structure their syllabi.[32] In other words, these differences do not reflect which instructors teach a course in a given academic year or how they structure their syllabi.

Taken together, these results show that AI is entering the classroom through early-career instructors who remain actively engaged in research. At the same time, neither the time available for teaching nor attention to AI in the syllabus explains which instructors integrate AI into their teaching. The arrival of generative AI thus reveals the diffusion of a new technology in real time, as these professors are the first to bring it into classrooms where it barely existed three years ago.

6.3 Heterogeneity Across Universities and Majors

The results presented above show that assistant professors with active research agendas are the instructors most likely to integrate AI into the classroom, yet the extent of AI integration varies across disciplinary and institutional contexts. I thus split the sample into STEM and non-STEM fields and re-estimate equation (7) for teaching integration within each group. Figure 11 shows that the main results hold in both groups. In particular, as-

[31] Whether the course explicitly allows students to use AI follows similar patterns, though instructors with publications, and particularly those who published recently, are also more likely to permit it (Figure A11).

[32] While these text-based characteristics are themselves outcomes of the syllabus-writing process, I include them to assess whether the instructor coefficients reflect genuine differences in AI adoption rather than differences in writing style.

sistant professors and instructors who published recently remain more likely to integrate AI into their teaching in STEM and non-STEM fields alike, though these estimates are less precise within each group, particularly in non-STEM fields.

Since AI has been differentially integrated into the classroom at selective institutions compared with the rest of the sample, I further examine whether these results differ across selectivity levels. The career-stage gap is larger at selective universities, where assistant professors are 1.04 percentage points more likely than full professors to integrate AI into their teaching. However, this gap does not differ significantly from the one at other universities. Research activity, on the other hand, plays a similar role across the selectivity spectrum, as the estimates for the years since the last publication are close in both groups. As such, recent research activity shapes AI integration at selective and less selective universities alike, while the career-stage gap is largest at selective universities.

Altogether, these results show that young, research-active faculty members lead the integration of AI in the classroom across fields and across the selectivity spectrum. These results thus align with growing evidence showing that AI adoption is more prevalent in firms with better organizational resources and more educated workers (McElheran et al., 2024; Bick et al., 2026). Early AI integration in higher education follows the same broad patterns as in the rest of the economy.

7 Conclusion

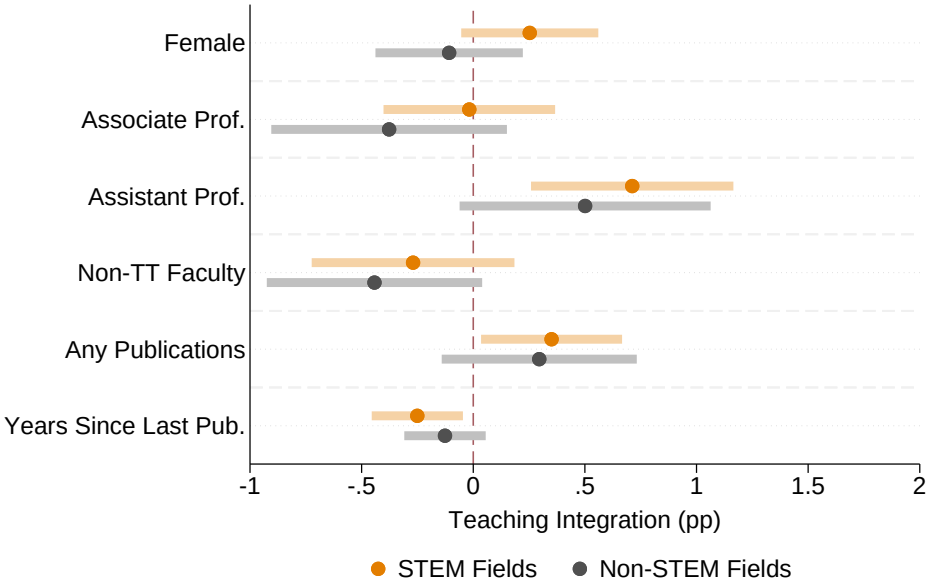
In this paper, I have examined how AI is being integrated in higher education, using a dataset of 327,495 course syllabi across 50 public universities linked to detailed instructor, course, and institutional characteristics. I have shown that explicit AI policy language has diffused widely, appearing in close to 73 percent of Fall 2026 syllabi. Yet this diffusion has not translated into comparable changes to what instructors teach or how they assess students, and only about six percent of syllabi show teaching integration.

To understand what drives these patterns, I have decomposed the variation in AI outcomes across instructors, courses, and institutions, and examined which observable instructor characteristics predict classroom integration. I find that instructors and the specific courses they teach account for the large majority of explained variation, while the university component is negligible. Institutions shape the conditions for formal policy diffusion, but they do not determine whether instructors actually redesign their courses around AI. Moreover, assistant professors and research-active instructors are the most likely to adopt AI in the classroom, and they do so primarily on the teaching margin rather than through evaluation. What matters is whether instructors are currently engaged in research, rather than how much they publish. Altogether, young, research-active faculty at the most research-intensive universities are leading the diffusion of AI into the classroom.

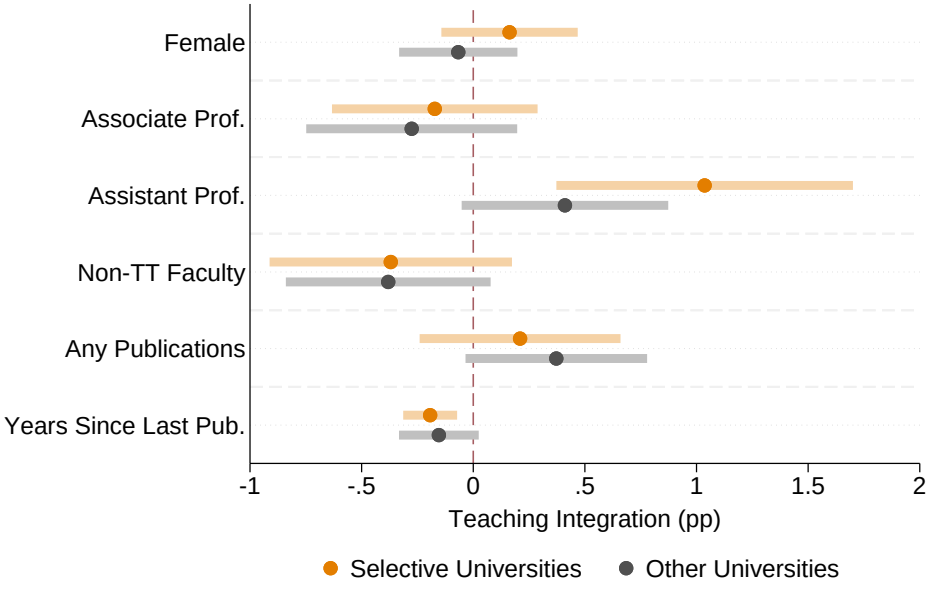
In this context, future work could link the syllabus-based measures developed here to student-level administrative records to estimate the returns to learning from a professor who was an early adopter of AI in the classroom (Biasi and Ma, 2026). Whether these

Figure 11: Heterogeneity in Teaching Integration

(a) STEM and Non-STEM Fields



(b) Selective and Other Universities



Note: Figure 11 presents coefficient estimates from regressions of teaching integration, measured in percentage points, on instructor characteristics entered jointly. Panel (a) estimates the specification separately within the STEM broad major group and all other majors. Panel (b) estimates it separately within the nine selective universities, whose 2019 entering class had a mean SAT score of 1250 or more, and the other 41 universities. Full professors are the omitted rank. All specifications absorb course and university-by-major-by-level-by-term fixed effects. Standard errors are clustered at the university level. The sample covers the 50 universities from Fall 2024 through Fall 2026. Bars represent 90% confidence intervals.

patterns persist as the technology matures, or whether they spread more broadly across institutions and fields, will determine how AI reshapes higher education in the years ahead.

References

Abel, J. R., Deitz, R., and Su, Y. (2014). Are recent college graduates finding good jobs? *Current Issues in Economics and Finance*, 20(1):1–8.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). AI and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.

Ash, E., Calderon, S., Chatzivasileiadis, M., and Hansen, S. (2025). Using LLMs to analyze interview transcripts at scale. Working paper, ETH Zürich / University College London.

Ash, E. and Hansen, S. (2023). Text algorithms in economics. *Annual Review of Economics*, 15:659–688.

Audoly, R., McGee, R., Ocampo, S., and Paz-Pardo, G. (2025). A practitioner's note on the Shapley-Owen-Shorrocks decomposition. Staff Reports 1163, Federal Reserve Bank of New York.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2024). Firm investments in artificial intelligence technologies and changes in workforce composition. Working Paper 31325, NBER.

Bacher-Hicks, A., Kane, T. J., and Staiger, D. O. (2014). Validating teacher effect estimates using changes in teacher assignments in Los Angeles. Working Paper 20657, National Bureau of Economic Research.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636.

Barrie, C., Palmer, A., and Spirling, A. (2024). Replication for language models: Problems, principles, and best practice for political science. Working Paper.

Bartik, A. W., Gupta, A., and Milo, D. (2024). The costs of housing regulation: Evidence from generative regulatory measurement. Working paper.

Biasi, B. and Ma, S. (2022). The education-innovation gap. Working Paper 29853, NBER.

Biasi, B. and Ma, S. (2026). Frontier knowledge in college and student success. Working paper, Yale SOM / NBER affiliation version. February 26, 2026 version.

Bick, A., Blandin, A., and Deming, D. J. (2025). The rapid adoption of generative AI. Working Paper 32337, NBER.

Bick, A., Blandin, A., Deming, D. J., Fuchs-Schündeln, N., and Jessen, J. (2026). Mind the gap: AI adoption in Europe and the U.S. Working Paper 34995, National Bureau of Economic Research.

Bonney, K., Breaux, C., Buffington, C., Dinlersoz, E., Foster, L., Goldschlag, N., Haltiwanger, J., Kroff, Z., and Savage, K. (2024). Tracking firm use of AI in real time: A snapshot from the business trends and outlook survey. U.S. Census Bureau working paper / report.

Braga, M., Paccagnella, M., and Pellizzari, M. (2016). The impact of college teaching on students' academic and labor market outcomes. *Journal of Labor Economics*, 34(3):781–822.

Braghieri, L., Eichmeyer, S., Levy, R., Mobius, M., Steinhardt, J., and Zhong, R. (2024). Article-level slant and polarization of news consumption on social media. CEPR Discussion Paper 19807.

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies: "engines of growth"? *Journal of Econometrics*, 65(1):83–108.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2025). Generative AI at work. *Quarterly Journal of Economics*, 140(2):889–942.

Bursztyn, L., Handel, B. R., Jiménez-Durán, R., and Roth, C. (2025). When product markets become collective traps: The case of social media. *American Economic Review*, 115(12):4105–4136.

Calderon, N., Reichart, R., and Dror, R. (2025). The alternative annotator test for LLM-as-a-judge: How to statistically justify replacing human annotators with LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 16051–16081, Vienna, Austria.

Carrell, S. E. and West, J. E. (2010). Does professor quality matter? evidence from random assignment of students to professors. *Journal of Political Economy*, 118(3):409–432.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the united states. *Quarterly Journal of Economics*, 135(3):1567–1633. Previously circulated as "Mobility Report Cards" (NBER Working Paper No. 23618).

Chiang, W.-L., Zheng, L., Sheng, Y., Angelopoulos, A. N., Li, T., Li, D., Zhang, H., Zhu, B., Jordan, M., Gonzalez, J. E., and Stoica, I. (2024). Chatbot arena: An open platform for evaluating LLMs by human preference. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*.

Chirikov, I. (2026). How instructors regulate AI in college: Evidence from 31,000 course syllabi. Higher Education Working Paper 26-1, UC Berkeley Center for Studies in Higher Education.

Chyn, E., Haggag, K., and Maruthiah, C. (2025). Ideology in government: Evidence from the office of Indian Affairs and the assimilation era. Working Paper 34415, National Bureau of Economic Research.

Clayton, C., Coppola, A., Maggiori, M., and Schreger, J. (2025). Geoeconomic pressure. Working Paper 34020, National Bureau of Economic Research.

Comin, D. and Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, 100(5):2031–2059.

Contractor, Z. and Reyes, G. (2025). Generative AI in higher education: Evidence from an elite college. Working paper, Middlebury College.

De Vlieger, P., Jacob, B., and Stange, K. (2016). Measuring instructor effectiveness in higher education. Working Paper 22998, NBER.

Dell, M. (2025). Deep learning for economists. *Journal of Economic Literature*, 63(1):5–58.

Dillon, E., Guo, N., and Saurer, R. (2025). Shifting work patterns in the age of generative AI. Working Paper 33845, NBER.

Dolnicar, S., Grün, B., and Leisch, F. (2011). Quick, simple and reliable: Forced binary survey questions. *International Journal of Market Research*, 53(2):231–252.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of large language models. *Science*, 384(6702):1306–1308.

Feld, J., Salamanca, N., and Zölitz, U. (2020). Are professors worth it? the value-added and costs of tutorial instructors. *Journal of Human Resources*, 55(3):836–863.

Felten, E. W., Raj, M., and Seamans, R. (2023). How will language modelers like ChatGPT affect occupations and industries? arXiv preprint arXiv:2303.01157.

Figlio, D. N., Schapiro, M. O., and Soter, K. B. (2015). Are tenure track professors better teachers? *Review of Economics and Statistics*, 97(4):715–724.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6):1176–1209.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Haaland, I., Roth, C., Stantcheva, S., and Wohlfart, J. (2025). Understanding economic behavior using open-ended survey data. *Journal of Economic Literature*, 63(4):1244–1280.

Handa, K., Bent, D., Tamkin, A., McCain, M., Durmus, E., Stern, M., Schiraldi, M., Huang, S., Ritchie, S., Syverud, S., Jagadish, K., Vo, M., Bell, M., and Ganguli, D. (2025). How university students use Claude. Technical report, Anthropic.

Hartley, J., Jolevski, F., Melo, V., and Moore, B. (2025). The labor market effects of generative artificial intelligence. Working Paper 5136877, SSRN.

Hoffmann, F. and Oreopoulos, P. (2009). Professor qualities and student achievement. *Review of Economics and Statistics*, 91(1):83–92.

Humlum, A. and Vestergaard, E. (2025). Large language models, small labor market effects. Working Paper 32424, NBER.

Javadian Sabet, A., Bana, S. H., Yu, R., and Frank, M. R. (2024). Course-skill atlas: A national longitudinal dataset of skills taught in u.s. higher education curricula. *Scientific Data*, 11:1410.

Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Working Paper 14607, National Bureau of Economic Research.

Korinek, A. (2023). Generative AI for economic research: Use cases and implications for economists. *Journal of Economic Literature*, 61(4):1281–1317.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Sage Publications, Thousand Oaks, CA, 2nd edition.

Lagakos, D., Michalopoulos, S., and Voth, H.-J. (2025). American life histories. Working Paper 33373, National Bureau of Economic Research.

Link, S., Peichl, A., Roth, C., and Wohlfart, J. (2024). Attention to the macroeconomy. Working Paper 32988, NBER.

Lombard, M., Snyder-Duch, J., and Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4):587–604.

Ludwig, J., Mullainathan, S., and Rambachan, A. (2025). Large language models: An applied econometric framework. Working Paper 33344, National Bureau of Economic Research.

McDonald, N., Johri, A., Ali, A., and Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 3:100121.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., and Zolas, N. (2024). AI adoption in america: Who, what, and where. *Journal of Economics & Management Strategy*, 33(2). NBER Working Paper 31788.

O'Connor, C. and Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19:1–13.

Priem, J., Piwowar, H., and Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2):247–252.

Saltiel, F. (2026). Imagined futures and early retirement outcomes. Working Paper, McGill University.

Walters, C. R. (2024). Empirical Bayes methods in labor economics. In Dustmann, C. and Lemieux, T., editors, *Handbook of Labor Economics*, volume 5, pages 183–260. North-Holland/Elsevier.

Yotzov, I., Barrero, J. M., Bloom, N., and Davis, S. J. (2025). Firm data on AI. Working Paper 34836, NBER.

- SUPPLEMENTARY APPENDICES -
For Online Publication

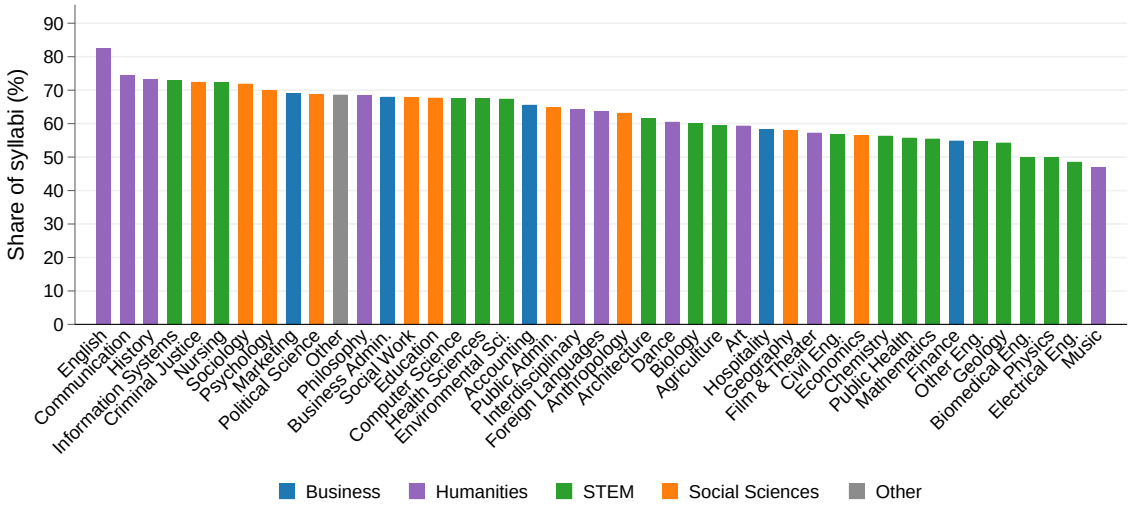
– **Appendix A**: Additional Figures and Tables ..... p. A2

– **Appendix B**: Data Construction ..... p. A26

– **Appendix C**: Classification Pipeline Details ..... p. A36

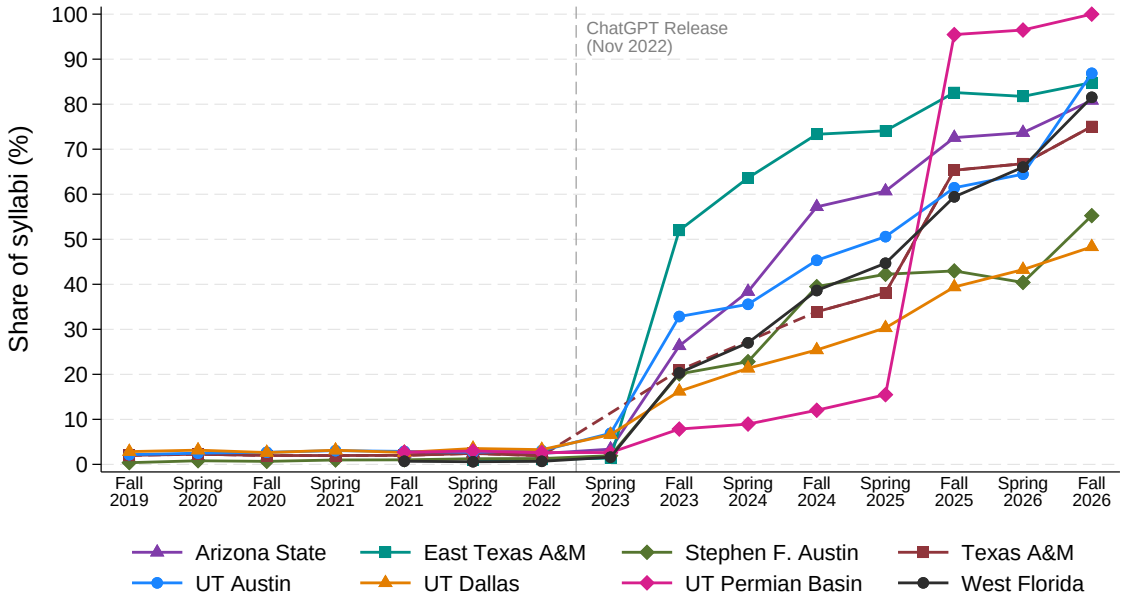
Appendix A: Additional Figures and Tables

Figure A1: AI Mentions Across Majors



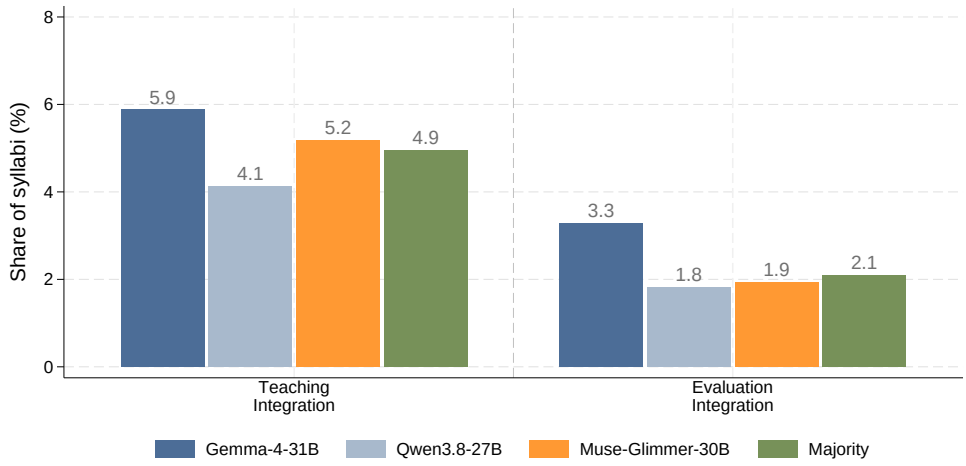
Note: Figure A1 shows the share of syllabi that mention artificial intelligence in each harmonized major, pooling the main sample from Fall 2024 through Fall 2026. Majors are ordered from the highest to the lowest share, and the colors mark the five macro-fields.

Figure A2: AI Mentions Over Time



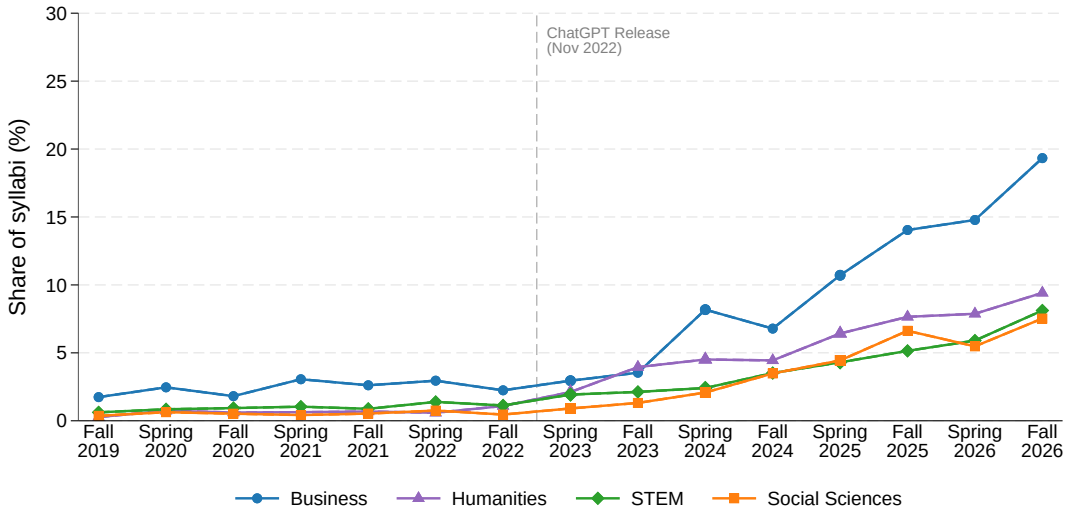
Note: Figure A2 shows the share of syllabi that mention artificial intelligence in each term from Fall 2019 through Fall 2026 at the eight universities in the longitudinal sample, one line per university. From Fall 2024 on, the figure uses the main-sample syllabi from the departments that each university offers in every earlier term. Texas A&M has no syllabi in Spring 2023 and Spring 2024, and a dashed segment bridges these terms. The vertical dashed line marks the public release of ChatGPT in November 2022.

Figure A3: Prevalence Across the Three Classification Models



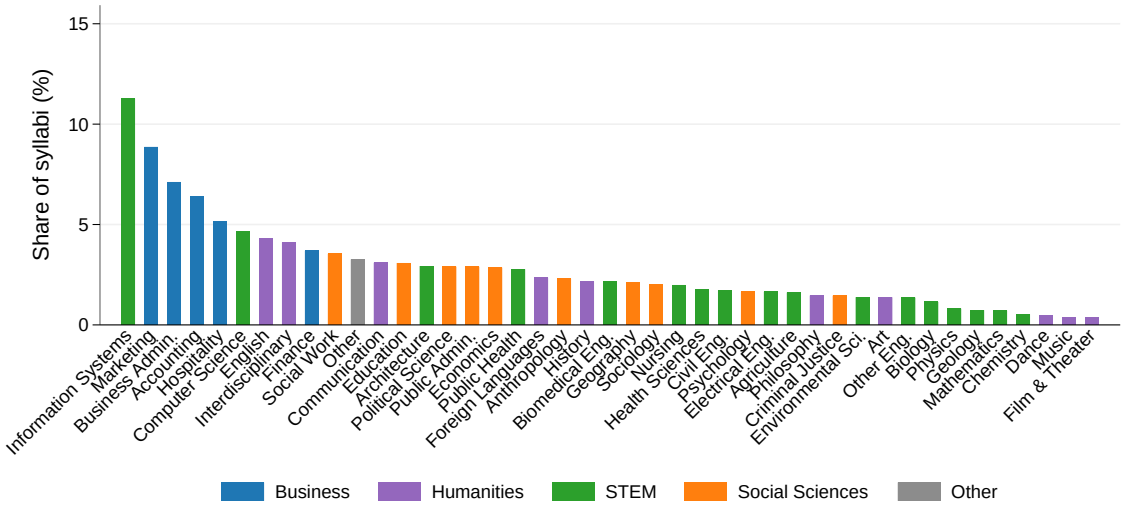
Note: Figure A3 shows the share of syllabi with Teaching Integration and Evaluation Integration under each of the three classification models, Gemma-4-31B, Qwen3.8-27B and Muse-Glimmer-30B, and under their majority, which is the measure used in the paper. Syllabi that never mention AI, or mention it only in the AI policy, are coded as not having either measure for every model. Each model is scored on the syllabi it rated, pooling all terms. AI Policy and AI Permitted come from a single classifier and are not shown.

Figure A4: Teaching Integration by Broad Field, Longitudinal



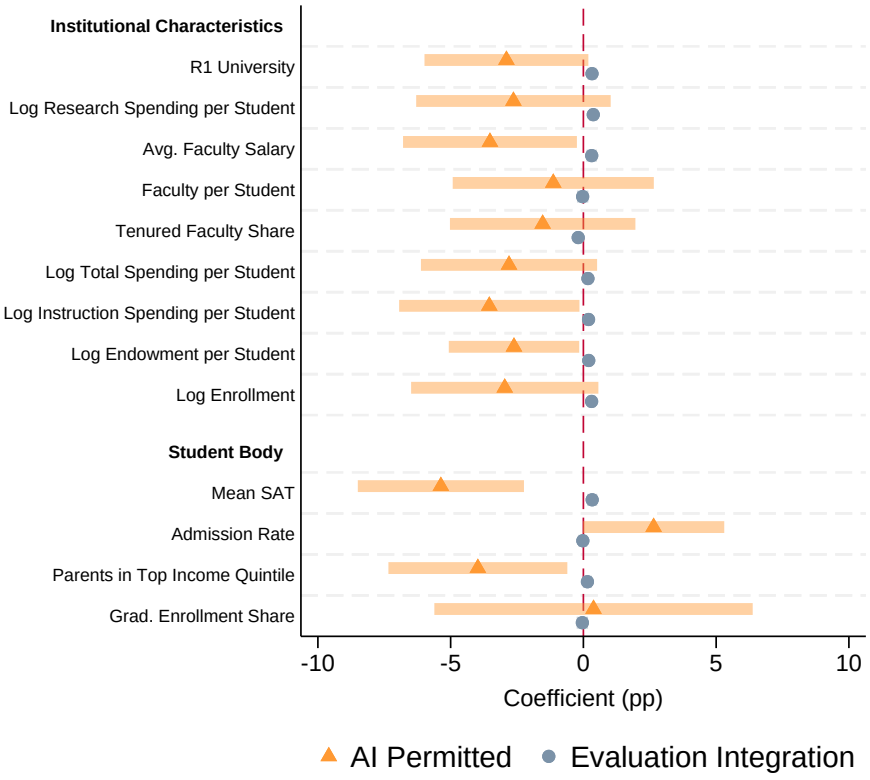
Note: Figure A4 shows the share of syllabi with Teaching Integration by broad field and semester from Fall 2019 through Fall 2026, at the four universities observed since Fall 2019 (Stephen F. Austin, Texas A&M, UT Austin, and UT Dallas). Each line gives the share at the universities observed in that semester. In the semesters in which one university is not observed (Stephen F. Austin in Spring 2025, and Texas A&M in Spring 2023 and Spring 2024), the share covers the other universities. The vertical dashed line marks the launch of ChatGPT in November 2022.

Figure A5: Evaluation Integration by Major



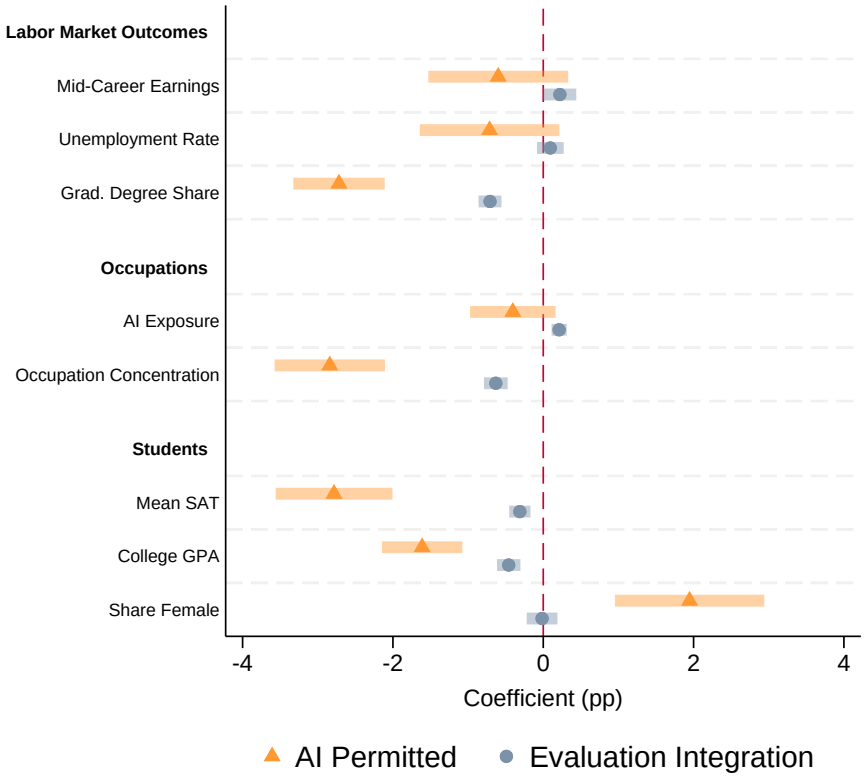
Note: Figure A5 presents the share of syllabi with AI evaluation integration across the harmonized majors in Fall 2026. Bars are colored by broad field (Business, Humanities, STEM, Social Sciences, Other).

Figure A6: University Characteristics: AI Permitted and Evaluation Integration



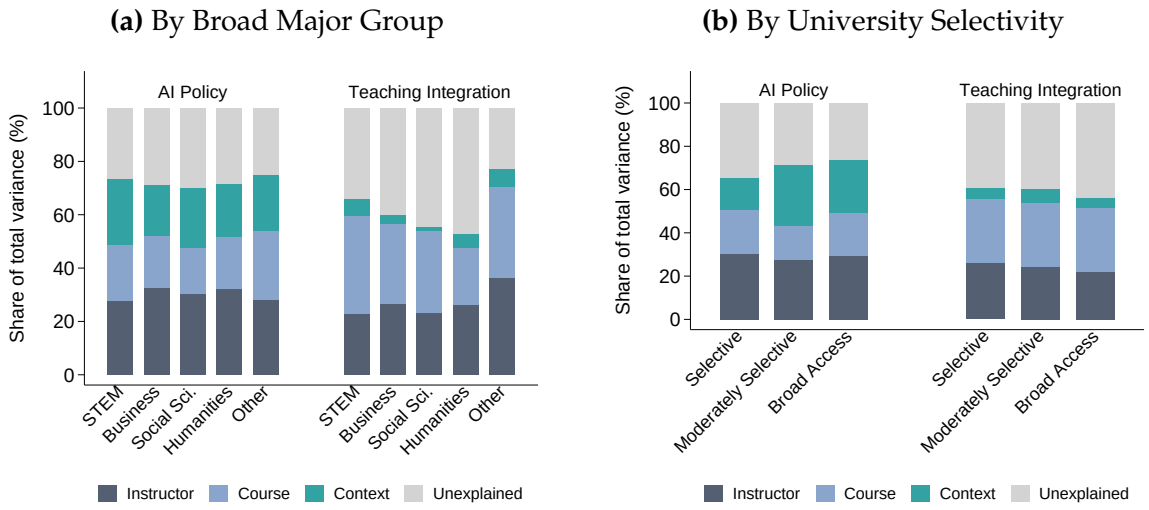
Note: Figure A6 presents coefficient estimates from regressions of AI Permitted and Evaluation Integration on one university characteristic at a time, pooling the five semesters from Fall 2024 to Fall 2026. Each row is a separate regression. I standardize each characteristic to mean zero and unit standard deviation across universities, each counted once. The outcome is in percentage points. Bars represent 90% confidence intervals. Fixed effects are major by course level by term. Standard errors are clustered at the university level. I set missing characteristic values to the mean and include a missing-value indicator.

Figure A7: Major Characteristics: AI Permitted and Evaluation Integration



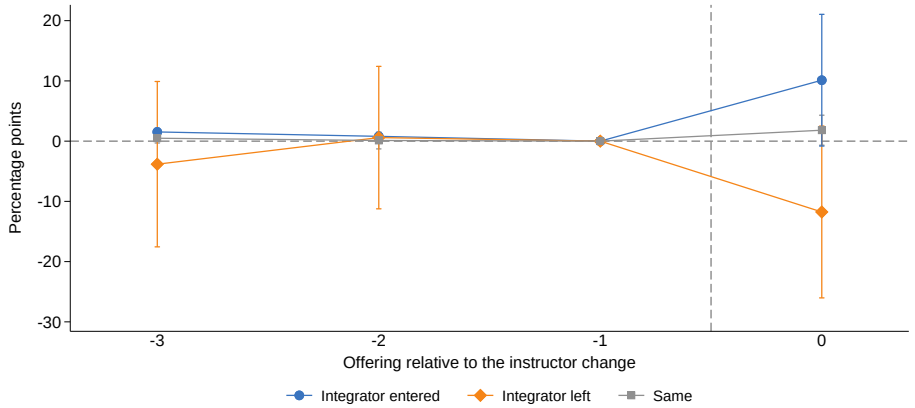
Note: Figure A7 presents coefficient estimates from regressions of AI Permitted and Evaluation Integration on one major characteristic at a time, pooling the five semesters from Fall 2024 to Fall 2026. Each row is a separate regression. I standardize each characteristic to mean zero and unit standard deviation across the harmonized majors, each counted once. The outcome is in percentage points. Bars represent 90% confidence intervals. Fixed effects are university and course level by term. Standard errors are clustered at the university level. I set missing characteristic values to the mean and include a missing-value indicator.

Figure A8: Variance Decomposition by Broad Field and Selectivity



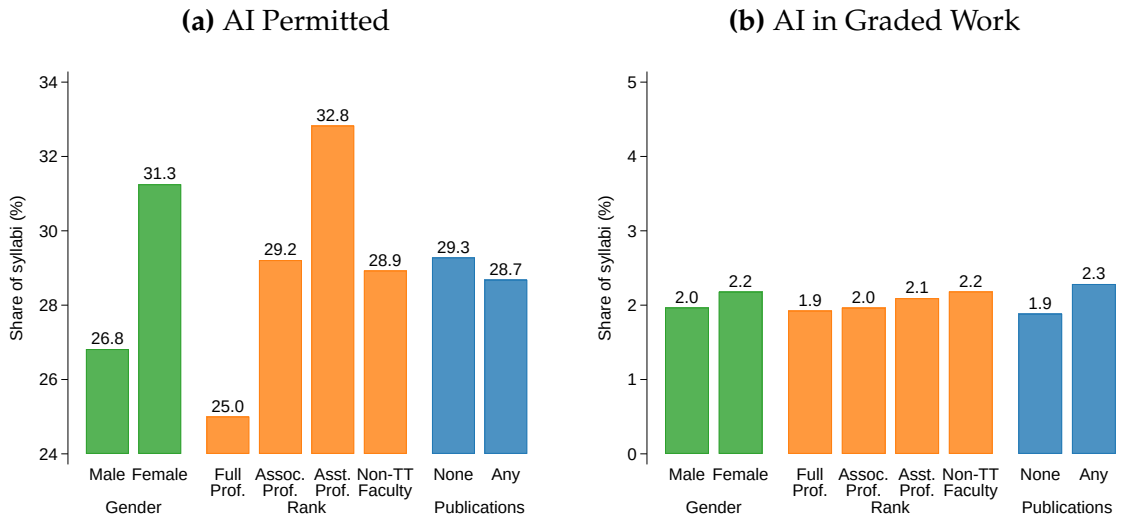
Note: Figure A8 shows the baseline three-way variance decomposition for AI Policy and Teaching Integration across subsamples. Panel (a) shows the decomposition separately across the five broad fields. Panel (b) shows the decomposition separately across the three selectivity tiers, for the universities with a selectivity rating. Each bar shows the shares of total variance attributed to instructor fixed effects, local course fixed effects, the broader context component, and the unexplained residual.

Figure A9: Event Study by Type of Instructor Change, Three Offerings Before the Change



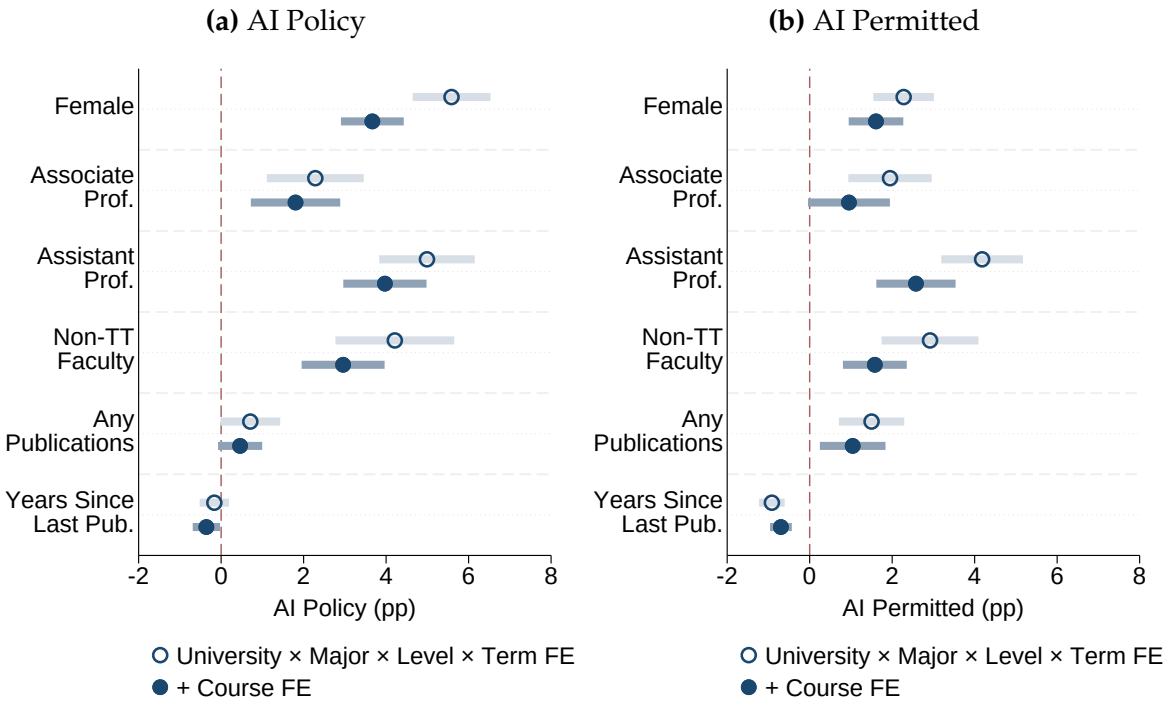
Note: Figure A9 presents estimates from an event-study version of equation (4) on courses observed in three offerings before the instructor change, with the last offering before the change as the reference, relative to courses at the same university whose instructor stays. Bars represent 95% confidence intervals based on standard errors clustered by university.

Figure A10: AI Permitted and AI in Graded Work by Instructor Characteristics



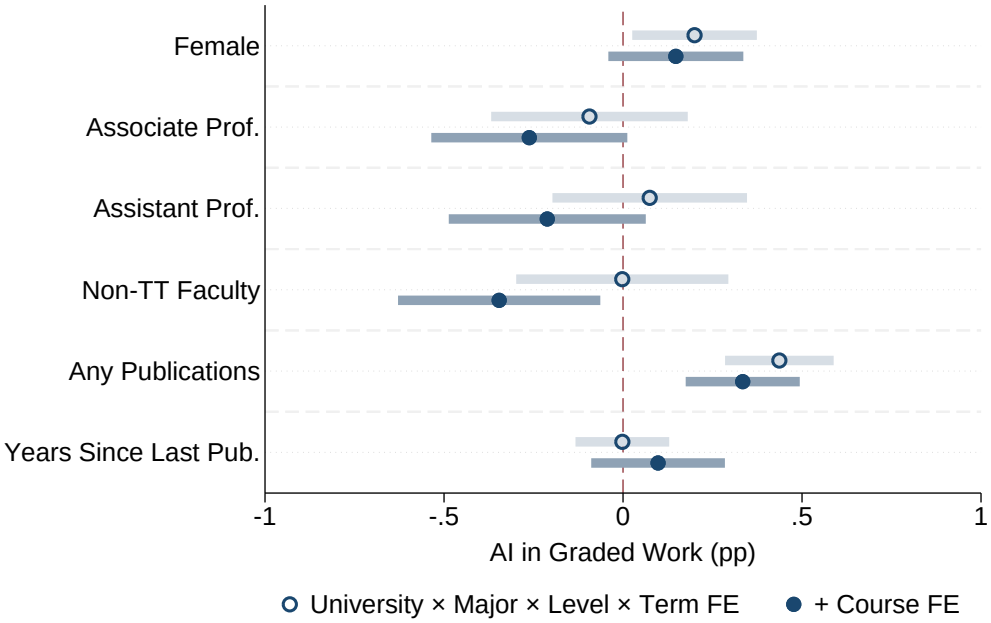
Note: Figure A10 presents the share of syllabi that explicitly permit students to use AI (panel a) and that integrate AI into graded work (panel b), in percent, by the gender, academic rank and publication record of the instructor. Each group includes the syllabi whose instructor has that characteristic recorded. The sample covers the 50 universities from Fall 2024 through Fall 2026.

Figure A11: Instructor Characteristics, AI Policy and AI Permitted



Note: Figure A11 presents coefficient estimates from regressions of whether the syllabus includes an explicit AI policy (panel a) and whether it explicitly permits students to use AI (panel b), measured in percentage points, on instructor characteristics entered jointly. Hollow circles show estimates from the specification with university-by-major-by-level-by-term fixed effects. Solid circles add course fixed effects. Both specifications are estimated on the same sample. Full professors are the omitted rank. Years Since Last Pub. is measured in standard deviations among instructors with publications and equals zero for instructors without publications. Standard errors are clustered at the university level. The sample covers the 50 universities from Fall 2024 through Fall 2026. Bars represent 90% confidence intervals.

Figure A12: Instructor Characteristics and AI in Graded Work



Note: Figure A12 presents coefficient estimates from regressions of whether the syllabus integrates AI into graded work, measured in percentage points, on instructor characteristics entered jointly. Hollow circles show estimates from the specification with university-by-major-by-level-by-term fixed effects. Solid circles add course fixed effects. Both specifications are estimated on the same sample. Full professors are the omitted rank. Years Since Last Pub. is measured in standard deviations among instructors with publications and equals zero for instructors without publications. Standard errors are clustered at the university level. The sample covers the 50 universities from Fall 2024 through Fall 2026. Bars represent 90% confidence intervals.

Table A1: Number of Syllabi by University and Academic Term

Panel A. Main Sample

University	Fall 2024	Spring 2025	Fall 2025	Spring 2026	Fall 2026	Total
Texas						
Texas A&M	3,422	3,443	4,319	4,960	5,816	21,960
UT Austin	2,944	3,133	3,153	2,986	3,264	15,480
Sam Houston State	1,948	1,947	2,239	2,303	2,803	11,240
North Texas	2,172	2,192	2,488	2,501	1,861	11,214
UT Dallas	2,055	2,058	2,163	2,104	2,146	10,526
UT San Antonio	2,041	2,076	2,073	2,017	2,261	10,468
Houston	0	0	3,342	3,263	3,254	9,859
UT Arlington	1,619	1,622	2,029	2,273	2,269	9,812
Stephen F. Austin	1,487	1,347	1,775	1,860	1,449	7,918
Texas Tech	1,567	1,403	1,636	1,560	1,538	7,704
Lamar	1,406	1,375	1,439	1,409	1,604	7,233
UT El Paso	1,340	1,220	1,406	1,585	1,432	6,983
East Texas A&M	1,163	1,137	1,255	1,435	1,212	6,202
Prairie View A&M	0	1,220	1,254	1,285	1,139	4,898
West Texas A&M	932	913	924	915	938	4,622
Texas A&M Kingsville	866	857	868	814	1,036	4,441
Tarleton State	0	0	1,341	1,319	1,623	4,283
Texas Southern	848	833	747	743	1,081	4,252
Texas A&M San Antonio	739	600	841	822	924	3,926
Houston Clear Lake	0	0	1,127	1,072	1,057	3,256
UT Permian Basin	526	576	627	605	604	2,938
Houston Downtown	0	0	0	1,413	1,380	2,793
Texas A&M International	0	0	874	832	789	2,495
Texas State	0	0	0	0	2,421	2,421
UT Tyler	0	0	573	590	1,042	2,205
Texas A&M Corpus Christi	0	0	0	827	1,250	2,077
Florida						
Central Florida	3,800	3,707	4,020	3,950	4,127	19,604
Florida International	0	0	3,640	4,020	4,425	12,085
Florida State	0	0	0	4,136	4,013	8,149
Florida Gulf Coast	1,514	1,399	1,567	1,749	1,868	8,097
West Florida	1,377	1,357	1,356	1,314	1,559	6,963
Florida	0	0	0	0	5,508	5,508
Florida A&M	0	0	1,313	1,344	1,517	4,174
North Florida	0	0	0	1,792	1,798	3,590
Georgia						
Georgia	1,721	1,708	2,272	3,076	3,242	12,019
Kennesaw State	0	0	0	5,423	4,535	9,958
Albany State	942	926	955	930	890	4,643
Georgia Tech	0	0	0	0	3,719	3,719
Georgia State	0	0	0	0	3,226	3,226
Georgia Southern	0	0	0	0	2,981	2,981
Augusta	0	0	0	0	2,156	2,156
North Georgia	0	0	0	0	2,002	2,002
Georgia College	0	0	0	783	1,003	1,786
West Georgia	0	0	0	0	1,610	1,610
Middle Georgia State	0	0	0	0	1,202	1,202
Clayton State	0	0	0	0	739	739
Indiana						
IU Bloomington	0	0	2,653	2,679	2,692	8,024
Other States						
Arizona State	3,918	3,766	3,590	3,771	2,783	17,828
Western Kentucky	0	0	1,666	1,555	1,661	4,882
North Carolina Central	0	0	0	0	1,344	1,344
Total	**40,347**	**40,815**	**61,525**	**78,015**	**106,793**	**327,495**

Note: Table A1 presents the number of syllabi in each term, separately by university. Panel A covers the main sample from Fall 2024 through Fall 2026, and Total sums its five terms. Universities are grouped by state, with Texas, Florida, Georgia and Indiana shown on their own and the remaining states pooled, and ordered within each group by total syllabi. Of the 50 universities, 23 are observed in every term, while all 50 contribute syllabi in Fall 2026. A zero means that no syllabi were collected for that university in that term.

Table A1: Number of Syllabi by University and Academic Term (Continued)

Panel B. Earlier Terms in the Longitudinal Sample

University	Fall 2019	Spring 2020	Fall 2020	Spring 2021	Fall 2021	Spring 2022	Fall 2022	Spring 2023	Fall 2023	Spring 2024	Total
UT Austin	2,819	2,809	2,894	2,889	3,017	2,944	2,944	3,044	2,908	2,990	29,258
Texas A&M	3,654	3,658	3,729	3,612	3,341	3,292	3,394	0	3,452	0	28,132
Arizona State	0	0	0	0	3,232	3,651	3,073	3,487	3,067	3,479	19,989
Stephen F. Austin	1,577	1,751	1,530	1,570	1,512	1,457	1,487	1,426	1,483	1,382	15,175
UT Dallas	1,023	1,107	1,103	1,112	1,056	1,078	1,131	1,166	1,145	1,176	11,097
West Florida	0	0	0	0	1,008	1,193	1,193	1,194	1,224	1,237	7,049
East Texas A&M	0	0	0	0	0	1,266	1,278	1,291	1,289	1,251	6,375
UT Permian Basin	0	0	0	0	590	507	535	493	510	469	3,104
Total	**9,073**	**9,325**	**9,256**	**9,183**	**13,756**	**15,388**	**15,035**	**12,101**	**15,078**	**11,984**	**120,179**

Note: Panel B of Table A1 presents the number of syllabi from Fall 2019 through Spring 2024 at the eight universities whose earlier terms I also collect. These syllabi are not part of the main sample. Together with the main-sample syllabi from the same universities, limited to the departments that each university offers in every earlier term, they form the longitudinal sample. A zero means that no syllabi were collected for that university in that term, and the Texas A&M syllabi for Spring 2023 and Spring 2024 are not available.

Table A2: Instructor Characteristics by University

University	Instructors (1)	Instructor Coverage (%) Gender (2)	Rank (3)	Any Pub. (4)
Texas				
East Texas A&M	1,230	89	67	22
Houston	2,549	94	81	55
Houston Clear Lake	626	96	59	38
Houston Downtown	772	97	89	37
Lamar	845	94	68	38
North Texas	2,312	94	84	46
Prairie View A&M	813	97	94	45
Sam Houston State	2,022	91	66	41
Stephen F. Austin	964	94	73	37
Tarleton State	907	93	88	30
Texas A&M	4,717	91	75	65
Texas A&M Corpus Christi	717	91	75	45
Texas A&M International	538	90	91	38
Texas A&M Kingsville	612	87	77	43
Texas A&M San Antonio	524	91	90	32
Texas Southern	689	85	33	32
Texas State	1,538	91	73	39
Texas Tech	2,096	88	70	53
UT Arlington	1,585	91	89	55
UT Austin	3,946	94	81	62
UT Dallas	2,366	80	67	47
UT El Paso	1,228	95	89	51
UT Permian Basin	455	95	67	33
UT San Antonio	1,902	92	81	54
UT Tyler	600	87	48	26
West Texas A&M	631	92	85	43
Florida				
Central Florida	3,326	90	74	54
Florida	3,184	87	60	62
Florida A&M	667	86	82	48
Florida Gulf Coast	1,325	93	60	35
Florida International	2,879	90	74	47
Florida State	3,065	89	60	52
North Florida	1,262	93	82	45
West Florida	1,077	96	45	32
Georgia				
Albany State	415	87	80	31
Augusta	886	87	83	34
Clayton State	274	91	73	22
Georgia	3,378	88	65	50
Georgia College	483	95	84	52
Georgia Southern	1,216	90	79	50
Georgia State	1,795	90	81	48
Georgia Tech	1,549	81	66	68
Kennesaw State	2,570	88	74	39
Middle Georgia State	451	93	78	10
North Georgia	839	91	79	28
West Georgia	870	92	73	23
Indiana				
IU Bloomington	3,354	77	62	47
Other States				
Arizona State	4,420	88	68	47
North Carolina Central	722	83	81	26
Western Kentucky	1,192	90	47	37
All 50 universities	**78,383**	**90**	**72**	**47**

Note: Table A2 presents instructor coverage by university, pooling the main sample from Fall 2024 through Fall 2026. Instructors is the number of distinct instructors teaching at the university over this period. The Gender, Rank and Any Pub. columns give the share of distinct instructors, in percent, with a recorded gender, an academic rank on the four-group ladder, and at least one publication. Universities are grouped by state, with Texas, Florida, Georgia and Indiana shown on their own and the remaining states pooled, and ordered alphabetically within each group. The bottom row pools all distinct instructors, so it is not the average of the rows above.

Table A3: Descriptive Statistics for the Longitudinal Sample

	Share of Sample (1)	**Share Mentioning AI** (2)
Instructor Gender		
Female	45.0	9.5
Male	55.0	8.0
Instructor Characteristics		
Tenure-Track	49.8	8.8
Any Publications	51.9	8.8
Major Characteristics		
Business	8.9	11.0
STEM	40.7	7.2
Social Sciences	18.2	9.0
Humanities	29.5	9.7
Course Characteristics		
Undergraduate	83.1	8.5
Graduate	16.9	9.5
Academic Year		
2019–20	15.3	2.0
2020–21	15.3	2.2
2021–22	24.3	2.2
2022–23	22.6	3.0
2023–24	22.5	29.8
Observations	120,179	10,258

Note: Table A3 presents descriptive statistics for the syllabi from Fall 2019 through Spring 2024 at the eight universities in the longitudinal sample. These syllabi are not part of the main sample. Column (1) shows the share of syllabi in each row, and column (2) shows the share of syllabi in each row that mention artificial intelligence. For the gender and tenure-track rows, I focus on syllabi with non-missing instructor information. Business, STEM, Social Sciences and Humanities denote the harmonized macro-fields, and Undergraduate combines freshman through senior courses. Arizona State, UT Permian Basin and West Florida enter the sample in Fall 2021 and East Texas A&M in Spring 2022, while Texas A&M has no syllabi in Spring 2023 and Spring 2024. The last row reports the number of syllabi in the sample and the number that mention artificial intelligence.

Table A4: Prevalence of the Four Measures by Semester

	Fall 2024	Spring 2025	Fall 2025	Spring 2026	Fall 2026
Policy					
AI Policy	41.5	45.3	64.3	68.4	82.6
AI Permitted	15.1	17.2	31.9	34.9	45.1
Classroom Integration					
Teaching Integration	3.1	3.7	5.3	5.5	7.0
Evaluation Integration	1.2	1.4	2.3	2.2	3.0
Syllabi	40,347	39,595	43,742	45,682	46,707
Institutions	23	23	23	23	23

Note: Table A4 presents the prevalence of the four measures semester by semester. Each entry gives the share of syllabi exhibiting the row measure, in percent. The table is restricted to the universities observed in all terms, so that the change across semesters is not driven by which campuses enter the data in later terms.

Table A5: AI Prevalence by Broad Major Field

	Business (1)	Humanities (2)	STEM (3)	Social Sci. (4)	Other (5)
AI Mentioned	65.2	65.7	60.3	67.6	68.6
Policy					
AI Policy	60.8	63.9	58.1	66.9	67.1
AI Permitted	34.3	27.0	27.6	31.8	27.7
Classroom Integration					
Teaching Integration	8.7	5.4	4.5	3.4	5.8
Evaluation Integration	4.5	2.1	1.6	1.9	2.2
Number of Syllabi	31,347	93,077	128,805	67,955	6,155

Note: Table A5 presents the share of syllabi with each measure, in percent, across the five broad fields into which harmonized majors are grouped, pooling all terms from Fall 2024 through Fall 2026. AI Mentioned is a keyword search over the full syllabus text. AI Policy and AI Permitted describe what the syllabus says about student use of AI. Teaching Integration and Evaluation Integration describe AI used in course activities and in graded work.

Table A6: AI Prevalence by Institution, Fall 2026

Institution	**Syllabi** (1)	**AI Policy** (2)	**AI Permitted** (3)	**Teaching Integration** (4)	**Evaluation Integration** (5)
Panel A. Selective					
Central Florida	4,127	99	59	6.9	2.8
Florida	5,508	35	20	7.8	3.7
Florida State	4,013	71	37	5.1	1.5
Georgia	3,242	82	37	6.4	2.3
Georgia Tech	3,719	34	22	3.9	1.6
IU Bloomington	2,692	39	18	6.1	2.1
Texas A&M	5,816	72	20	9.6	2.9
UT Austin	3,264	86	52	10.4	3.4
UT Dallas	2,146	45	26	13.7	6.1
Tier total and mean	*34,527*	*63*	*32*	*7.8*	*2.9*
Panel B. Moderately Selective					
Arizona State	2,783	81	42	7.4	3.5
Augusta	2,156	51	32	1.3	0.7
Florida Gulf Coast	1,868	76	37	4.3	3.6
Florida International	4,425	41	21	9.8	4.9
Georgia College	1,003	100	2	1.0	1.5
Georgia Southern	2,981	100	12	0.9	0.7
Georgia State	3,226	59	24	5.4	3.3
Houston	3,254	52	28	8.0	3.1
Houston Clear Lake	1,057	63	29	4.3	1.5
Kennesaw State	4,535	93	39	5.7	2.3
North Florida	1,798	69	37	5.3	2.7
North Texas	1,861	70	27	6.3	2.1
Texas A&M Corpus Christi	1,250	86	49	4.4	1.4
Texas Tech	1,538	76	35	4.8	2.6
UT Arlington	2,269	79	42	8.0	4.0
UT San Antonio	2,261	97	59	8.2	3.9
UT Tyler	1,042	81	52	4.5	2.2
West Florida	1,559	82	51	3.1	1.7
Western Kentucky	1,661	65	27	2.7	1.6
Tier total and mean	*42,527*	*75*	*34*	*5.0*	*2.5*
Panel C. Broad Access					
Albany State	890	100	100	3.4	1.2
Clayton State	739	100	44	3.9	3.1
East Texas A&M	1,212	86	12	3.6	1.6
Florida A&M	1,517	18	10	5.1	3.4
Houston Downtown	1,380	97	40	6.7	2.9
Lamar	1,604	100	97	5.0	3.3
Middle Georgia State	1,202	100	38	2.0	0.9
North Carolina Central	1,344	69	55	1.9	0.8
North Georgia	2,002	72	33	6.9	2.5
Prairie View A&M	1,139	100	92	4.7	3.8
Sam Houston State	2,803	95	45	5.4	3.0
Stephen F. Austin	1,449	56	30	3.5	2.1
Tarleton State	1,623	90	47	3.5	1.0
Texas A&M International	789	100	36	5.6	2.8
Texas A&M Kingsville	1,036	100	74	3.6	1.4
Texas A&M San Antonio	924	83	37	7.9	2.1
Texas Southern	1,081	98	78	4.3	2.2
Texas State	2,421	87	52	7.6	4.0
UT El Paso	1,432	66	40	6.3	2.4
UT Permian Basin	604	100	48	3.0	1.3
West Georgia	1,610	88	37	4.7	1.8
West Texas A&M	938	98	91	6.6	3.9
Tier total and mean	*29,739*	*86*	*52*	*4.8*	*2.3*
All institutions	**106,793**	**78**	**41**	**5.4**	**2.5**

Note: Table A6 presents Fall 2026 prevalence for each institution, grouped by institutional selectivity tier. Column (1) is the number of syllabi the institution contributes and columns (2) to (5) are shares in percent. In the tier rows and the final row, column (1) is a total and columns (2) to (5) are unweighted averages across institutions.

Table A7: AI Prevalence by Course Level

	All	Basic Undergrad.	Advanced Undergrad.	Graduate
	(1)	(2)	(3)	(4)
Panel A. Fall 2024 to Fall 2026				
AI Mentioned	64.0	66.8	63.4	61.0
Policy				
AI Policy	62.0	65.7	61.3	58.0
AI Permitted	28.9	26.4	29.7	31.0
Classroom Integration				
Teaching Integration	4.9	3.9	4.8	6.8
Evaluation Integration	2.1	1.9	2.2	2.3
Number of Syllabi	327,495	103,466	153,688	70,154
Panel B. Fall 2026				
AI Mentioned	75.1	78.9	75.7	68.9
Policy				
AI Policy	73.2	77.7	73.6	66.2
AI Permitted	37.3	33.8	39.4	38.4
Classroom Integration				
Teaching Integration	6.1	5.1	5.8	7.8
Evaluation Integration	2.7	2.2	2.8	3.1
Number of Syllabi	106,793	34,730	46,745	25,311

Note: Table A7 presents the share of syllabi with each measure, in percent, by course level. Panel A pools all terms from Fall 2024 through Fall 2026 and Panel B is the Fall 2026 cross-section. Column (1) pools all syllabi, including those whose level could not be assigned, so it is not a weighted average of columns (2) to (4).

Table A8: Course Characteristics with Pre-2024 Course Information

	AI Policy (1)	AI Permitted (2)	Teaching Integration (3)	Evaluation Integration (4)
Panel A. Course Level				
Freshman	4.95***	0.32	-1.57**	-0.27
	(1.04)	(0.85)	(0.70)	(0.39)
Sophomore	4.62***	0.03	-2.40***	0.56
	(0.79)	(0.68)	(0.54)	(0.38)
Junior	3.50***	1.18**	-2.44***	0.24
	(0.69)	(0.59)	(0.45)	(0.29)
Senior	3.01***	1.48***	-1.34***	0.29
	(0.60)	(0.51)	(0.40)	(0.22)
Panel B. Other Course Characteristics				
Log Enrollment	1.00***	0.38	-0.33	0.02
	(0.38)	(0.32)	(0.28)	(0.13)
Has TA	3.10***	0.81	-0.36	-0.05
	(0.69)	(0.57)	(0.35)	(0.20)
Credit Hours	1.44***	1.04*	-0.01	-0.00
	(0.52)	(0.59)	(0.29)	(0.13)
Course Offered Before Fall 2024	2.16***	-1.10**	-1.11**	-0.71***
	(0.59)	(0.53)	(0.43)	(0.26)
Adjusted R^2	0.583	0.563	0.370	0.355
Observations	76,336	76,336	76,281	76,331

Note: Table A8 presents the multivariate course-characteristic coefficients for AI Policy, AI Permitted, Teaching Integration, and Evaluation Integration, restricted to the universities with earlier terms in Table A1, where pre-2024 course information is available. The specification adds an indicator for whether an earlier syllabus shows that the course was offered before Fall 2024 to the main course-characteristic bundle, while continuing to absorb instructor fixed effects and university-by-major-by-term fixed effects. Graduate courses are the omitted category and missing indicators are included for unavailable continuous characteristics and pre-2024 course status. Standard errors, two-way clustered by instructor and local course, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Variance Decomposition with Statewide Course Codes

	AI Policy (1)	AI Permitted (2)	Teaching Integration (3)	Evaluation Integration (4)
Instructor	27.1	31.0	26.1	24.9
Local Course	10.8	10.8	16.7	14.2
Statewide Course by Term	15.3	12.4	11.6	10.8
Context	16.2	10.4	7.7	4.7
Unexplained	30.5	35.4	37.9	45.3
Adjusted R^2	69.5	64.6	62.1	54.7
Observations	29,295	29,295	29,245	29,250
Universities	8	8	8	8

Note: Table A9 presents the harmonized-course decomposition for the four AI measures, estimated on the Florida universities, which share a common state course numbering system. The decomposition separates local course variation from statewide-course-by-term variation within a specification that also includes instructor fixed effects and the broader context component. Each entry gives the share of total outcome variance, in percentage points, attributed to the corresponding component.

Table A10: Integrators and Other AI Outcomes

	(1)	(2)	(3)	(4)
Panel A. AI Permitted				
Integrator Entered	7.61***	7.39***	7.26***	6.84*
	(1.23)	(1.60)	(1.59)	(4.00)
Integrator Left	-4.39*	-3.90*	-4.02*	-2.91
	(2.25)	(2.10)	(2.10)	(3.36)
Same	0.75	1.04**	0.95**	3.12*
	(0.48)	(0.46)	(0.47)	(1.54)
Mean Before the Change	23.87	23.87	23.87	29.90
Stayer Baseline	11.35	11.35	11.35	11.17
Changes	12,596	12,595	12,595	1,846
Stayer Pairs	42,397	42,395	42,395	15,058
Panel B. AI in Graded Work				
Integrator Entered	2.76***	2.61***	2.57***	3.44
	(0.75)	(0.67)	(0.66)	(2.29)
Integrator Left	-3.93***	-3.85***	-3.81***	-6.88***
	(0.72)	(0.69)	(0.69)	(2.07)
Same	-0.26	-0.27	-0.26	-0.37
	(0.17)	(0.20)	(0.20)	(0.38)
Mean Before the Change	1.54	1.54	1.54	1.83
Stayer Baseline	0.53	0.53	0.53	0.60
Changes	12,605	12,605	12,605	1,847
Stayer Pairs	42,433	42,433	42,433	15,087
University × Semester Pair FE	Yes	No	No	No
University × Field × Semester Pair FE	No	Yes	Yes	Yes
Instructor Rank	No	No	Yes	No
Seen Two Offerings Before	No	No	No	Yes
Universities	39	39	39	33

Note: Table A10 presents OLS estimates of the change in the share of a course's syllabi that explicitly permit students to use AI (Panel A) and that integrate AI into graded work (Panel B), in percentage points, from the last offering before an instructor change to the first offering after it. The instructor types, samples and specifications follow Table 4. An integrator is an instructor who integrated AI into teaching in at least one syllabus of another course taught before the change. The mean before the change is the average share at the first of the two offerings across all courses in the sample, and the stayer baseline is the mean change among courses whose instructor stays. Standard errors clustered by university in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A11: Publication and Citation Quartiles and Teaching Integration

	Publications		Citations	
	(1)	(2)	(3)	(4)
Bottom Quartile	0.37**	0.40*	0.33*	0.36*
	(0.18)	(0.21)	(0.19)	(0.20)
Second Quartile	0.64**	0.14	0.58**	0.12
	(0.27)	(0.23)	(0.26)	(0.23)
Third Quartile	1.13***	0.37	1.24***	0.40
	(0.33)	(0.32)	(0.28)	(0.32)
Top Quartile	1.33***	0.42*	1.30***	0.52**
	(0.26)	(0.24)	(0.23)	(0.26)
Equal Across Quartiles (p-value)	0.00	0.54	0.00	0.54
Outcome Mean	4.89	4.89	4.89	4.89
Course FE	No	Yes	No	Yes
University × Major × Level × Term FE	Yes	Yes	Yes	Yes
Observations	276,284	276,284	276,284	276,284
Universities	50	50	50	50

Note: Table A11 presents OLS regression estimates of teaching integration, in percentage points, on indicators for the quartile of the instructor's publication count (columns 1 and 2) or citation count (columns 3 and 4) within their field. Instructors without publications are the omitted group. All columns control for gender, academic rank and years since the last publication. Columns 1 and 3 absorb university-by-major-by-level-by-term fixed effects, and columns 2 and 4 add course fixed effects, on the same sample. The p-value tests that the four quartile coefficients are equal. The sample covers the 50 universities from Fall 2024 through Fall 2026. Standard errors clustered by university in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A12: Timing of the Last Publication and Teaching Integration

	(1)	(2)
Last Publication Before the Course Year		
0 to 1 Years	0.99***	0.49**
	(0.20)	(0.20)
2 to 3 Years	0.12	-0.07
	(0.27)	(0.22)
4 to 6 Years	-0.07	-0.14
	(0.32)	(0.28)
7 to 10 Years	0.64	0.14
	(0.45)	(0.40)
11 or More Years	-0.23	-0.14
	(0.27)	(0.26)
Outcome Mean	4.89	4.89
Course FE	No	Yes
University × Major × Level × Term FE	Yes	Yes
Observations	276,284	276,284
Universities	50	50

Note: Table A12 presents OLS regression estimates of teaching integration, in percentage points, on indicators for the number of years between the year the course is taught and the instructor's most recent publication. Instructors without publications are the omitted group. Both columns control for gender and academic rank. Column 1 absorbs university-by-major-by-level-by-term fixed effects, and column 2 adds course fixed effects, on the same sample. The sample covers the 50 universities from Fall 2024 through Fall 2026. Standard errors clustered by university in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A13: Instructor Characteristics and Teaching Integration: Robustness

	Main	Course × Academic Year	Text Controls
	(1)	(2)	(3)
Assistant Prof.	0.59**	0.67**	0.43*
	(0.23)	(0.27)	(0.22)
Associate Prof.	-0.23	-0.35	-0.33
	(0.20)	(0.24)	(0.20)
Non-TT Faculty	-0.37*	-0.27	-0.48**
	(0.20)	(0.23)	(0.19)
Any Publications	0.32*	0.30	0.30*
	(0.18)	(0.19)	(0.17)
Years Since Last Pub.	-0.18***	-0.18**	-0.19***
	(0.06)	(0.07)	(0.06)
Female	0.01	-0.10	-0.12
	(0.12)	(0.14)	(0.12)
Outcome Mean	4.89	4.84	4.89
Course FE	Yes	No	Yes
Course × Academic Year FE	No	Yes	No
University × Major × Level × Term FE	Yes	Yes	Yes
Text Controls	No	No	Yes
Observations	276,284	202,836	276,284
Universities	50	50	50

Note: Table A13 presents OLS regression estimates of teaching integration, in percentage points, on instructor characteristics. Full professors are the omitted rank. Column 1 repeats the main specification. Column 2 replaces course fixed effects with course-by-academic-year fixed effects, which excludes syllabi that are the only ones of their course in an academic year. Column 3 adds nine text controls: reading grade level, average sentence length, lexical diversity, the shares of prescriptive, penalty and supportive language, indicators for a late-work policy and an attendance policy, and syllabus length. The sample covers the 50 universities from Fall 2024 through Fall 2026. Standard errors clustered by university in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Appendix B: Data Construction

B.1 Main Sample

I collect course syllabi from 50 public universities across seven states. The sample covers 26 universities in Texas, eight in Florida, 12 in Georgia, and one each in Arizona, Indiana, Kentucky, and North Carolina. I draw on each university's publicly accessible syllabus records to collect documents for the academic terms between Fall 2024 and Fall 2026, with the set of terms observed varying across universities (Table A1). For eight of these universities, I additionally recover syllabi between Fall 2019 and Spring 2024, which provides a longer window for the longitudinal analyses.

Data sources. How universities make syllabi public varies across the 50 institutions in the sample. Some maintain dedicated PDF portals, while others use course management platforms with structured metadata or searchable systems organized by department and academic term. This variation determines how much course information I can recover directly from the source and how much I must extract from the syllabus text. Table B1 lists the universities in the sample and the number of syllabi I collect at each institution, together with characteristics that describe their size, selectivity, and research orientation.

University characteristics. To characterize the universities in the sample, I draw on IPEDS and the Opportunity Insights database of Chetty et al. (2020), which together cover enrollment, admissions selectivity, institutional resources, and students' socioeconomic composition. I use these characteristics when examining how the integration of AI varies across universities.

Course and instructor information. For each syllabus in the sample, I recover six core metadata fields, covering the instructor's name, the course title, the department code and course number, the section identifier, and the academic term in which the course is offered. These fields determine how each syllabus is linked to a specific instructor, course, and academic unit.

Three of these fields, the department code, course number, and academic term, are typically recorded directly by each university's data system rather than extracted from the syllabus text itself. I recover these for effectively all syllabi in the sample, as they are encoded in the file structure or in structured metadata fields provided by each university's portal. For the small number of records where the academic term is not available from the source, I recover it by searching the syllabus text for explicit references to the semester and academic year.

Instructor names and course titles require a more involved extraction approach because their format and placement vary across institutions. For instructor names, I search the opening pages of each syllabus using text-matching rules ordered by precision, beginning with

Table B1: Universities in the Sample

University	Total Enrollment	Mean SAT	Research Exp. per FTE ($)	Graduate Share (%)	Syllabi
Texas (26 universities)					
East Texas A&M	14,047	1,055	487	36	6,202
Houston	52,007	1,240	3,925	19	9,859
Houston Clear Lake	10,463	1,080	264	27	3,256
Houston Downtown	17,295	1,035	126	10	2,793
Lamar	21,648	1,020	416	56	7,233
North Texas	52,260	1,105	2,064	28	11,214
Prairie View A&M	10,095	940	2,705	13	4,898
Sam Houston State	23,907	1,040	590	16	11,240
Stephen F. Austin	12,740	1,070	316	14	7,918
Tarleton State	17,448	1,055	1,512	16	4,283
Texas A&M	78,501	1,260	13,322	23	21,960
Texas A&M Corpus Christi	13,671	1,075	2,979	34	2,077
Texas A&M International	9,968	1,025	554	21	2,495
Texas A&M Kingsville	7,003	1,070	3,571	23	4,441
Texas A&M San Antonio	8,479	960	526	10	3,926
Texas Southern	9,571	900	701	21	4,252
Texas State	41,676	1,085	3,256	12	2,421
Texas Tech	43,357	1,180	5,565	21	7,704
UT Arlington	54,487	1,110	2,391	31	9,812
UT Austin	55,710	1,360	17,781	20	15,480
UT Dallas	34,317	1,290	4,141	32	10,526
UT El Paso	28,323	980	4,815	17	6,983
UT Permian Basin	7,456	1,010	634	24	2,938
UT San Antonio	38,728	1,100	4,342	13	10,468
UT Tyler	11,335	1,115	2,926	27	2,205
West Texas A&M	10,667	1,060	834	27	4,622
Florida (8 universities)					
Central Florida	80,332	1,270	2,472	15	19,604
Florida	62,811	1,390	16,135	38	5,508
Florida A&M	10,428	1,090	3,138	16	4,174
Florida Gulf Coast	18,327	1,125	513	13	8,097
Florida International	69,591	1,155	3,118	18	12,085
Florida State	49,885	1,315	5,398	27	8,149
North Florida	19,673	1,100	928	14	3,590
West Florida	17,296	1,125	910	34	6,963
Georgia (12 universities)					
Albany State	7,603	820	212	8	4,643
Augusta	10,199	1,113	10,924	—	2,156
Clayton State	7,169	973	117	—	739
Georgia	44,713	1,275	12,019	25	12,019
Georgia College	7,397	1,150	132	20	1,786
Georgia Southern	29,617	1,070	1,384	15	2,981
Georgia State	42,129	1,160	5,516	21	3,226
Georgia Tech	53,737	1,447	25,372	—	3,719
Kennesaw State	49,914	1,120	360	11	9,958
Middle Georgia State	9,626	1,055	15	—	1,202
North Georgia	21,038	1,136	80	—	2,002
West Georgia	14,712	1,044	179	—	1,610
Indiana (1 university)					
IU Bloomington	61,491	1,285	2,376	19	8,024
Other States (3 universities)					
Arizona State	87,154	1,210	5,827	19	17,828
North Carolina Central	8,456	1,023	1,974	—	1,344
Western Kentucky	19,863	915	412	13	4,882
Total (50)					**327,495**

Note: Table B1 presents characteristics and syllabi counts for each university. I obtain total enrollment, mean SAT, research expenditures per full-time-equivalent student, and the graduate enrollment share from IPEDS. I compute mean SAT as the sum of the verbal and math section midpoints between the 25th and 75th percentile scores. A dash marks a value that is not observed.

explicitly labeled fields and then moving to less structured patterns such as academic honorifics and university email addresses. I then standardize each extracted name so that the same instructor is identified consistently across documents. For course titles, I apply an analogous approach, supplemented by institution-specific rules where formatting differs substantially.

For the remaining syllabi where the text-based extraction does not recover the instructor name or course title, I send the opening pages of the document to a large language model that returns the missing field along with a reliability indicator. I use a standardized prompt, reported in the replication materials. Altogether, I recover instructor names and course titles for 99 percent of syllabi across the 50 universities.

Sample construction. I restrict the sample to the academic terms between Fall 2024 and Fall 2026. I then apply quality filters that exclude records without substantive syllabus content, covering course listing pages, preliminary course descriptions, empty course management system shells, and documents too short to describe the course and its policies. I additionally exclude course types that do not correspond to classroom instruction, such as practicum and independent-study registrations, as well as first-year-experience courses.

Because instructors often teach several sections of the same course with identical syllabi, I deduplicate the syllabi at the section level, following Biasi and Ma (2022). I group syllabi by institution, department, course number, academic term, and normalized instructor name, and keep the longest document in each group. As such, the final estimation sample contains 327,495 course-section observations, covering 107,705 courses taught by 78,383 instructors across the 50 universities and 42 harmonized college majors.

B.2 Instructor Data

I build the instructor dataset from the syllabus metadata, which identify 78,383 unique instructors active between Fall 2024 and Fall 2026 across the 50 universities. I then link this roster to six external source families covering bibliometric records, education records, publicly available CVs and faculty profiles, department websites, state salary databases, and residual text-based or name-based sources. Across these sources, I recover academic rank for 72 percent of instructors, gender for 90 percent, and research indicators for 47 percent. Table A2 shows coverage across the 50 universities.

OpenAlex. I use OpenAlex, an open bibliometric database covering publications, citations, and author records across all academic fields, to recover measures of instructors' research productivity and career trajectories. I match the instructor roster to the pool of authors affiliated with each university, using exact and progressively more flexible name comparisons and reviewing ambiguous cases by hand. Match rates remain below one hundred percent because many instructors in the sample do not maintain indexed research profiles.

From matched OpenAlex profiles, I recover the h-index, publication count, citations, re-

search topics, and the years since first and last publication (Table B2). The first-publication year serves as a lower-bound career-stage proxy when direct PhD information is unavailable.

ORCID. Some instructors in the roster have ORCID identifiers linked through their OpenAlex profiles. I query these ORCID records, and supplement them when needed with name-based searches using institutional affiliation filters, to recover education records and related employment information. ORCID education records are the highest-priority source for PhD information because they are researcher verified, and they account for the largest share of the PhD completion years I recover across sources.

Instructors' CVs. I also recover instructors' educational backgrounds from publicly available CVs and faculty profile pages. I search the text of these documents for degree information and recover the year in which the instructor completed the PhD. These records provide PhD completion years for instructors who do not appear in ORCID, and they also help confirm academic rank where I do not observe titles elsewhere.

Faculty directories. I collect faculty directory information by downloading profile pages from departmental and institutional websites, from which I recover academic rank, PhD year, PhD institution, and research area. Because departments update these pages as faculty are hired and promoted, the recovered titles generally reflect instructors' current positions. Faculty directories are especially useful where salary data provide limited information on academic rank, particularly in Indiana.

State salary records. In this context, publicly available state salary records provide a further source of instructors' official titles. Because these records report the title under which each instructor is employed, they provide especially reliable information on academic rank. I match instructors to these records by name and normalize the reported job titles to a common four-level rank hierarchy, and Texas records additionally provide gender and hire date.

Other sources. For instructors whose academic rank I do not recover from salary databases or faculty directories, I extract rank information directly from syllabus text by searching instructor information blocks and signature lines for academic titles. For instructors without direct gender records, I impute gender using U.S. Social Security Administration first-name frequencies, assigning gender only when one category accounts for at least 75 percent of births associated with that first name.

Consolidating sources. Because several data sources may provide the same characteristic for a given instructor, I prioritize official administrative records over self-reported data, and self-reported data over web-based or syllabus-derived information. Applying this ordering, I classify instructors into four rank categories and recover gender, while OpenAlex

Table B2: Research Productivity among Matched Instructors by University

University	Instructors	Matched	Conditional on Match	
		(%)	Publications	Citations
Texas				
East Texas A&M	1,230	22.1	43	857
Houston	2,549	55.3	68	1,864
Houston Clear Lake	626	38.5	36	795
Houston Downtown	772	36.8	39	869
Lamar	845	37.9	57	970
North Texas	2,312	46.5	59	1,324
Prairie View A&M	813	45.3	48	920
Sam Houston State	2,022	40.5	32	597
Stephen F. Austin	964	36.5	23	290
Tarleton State	907	30.4	26	538
Texas A&M	4,717	65.3	87	2,417
Texas A&M Corpus Christi	717	45.2	39	787
Texas A&M International	538	37.5	31	590
Texas A&M Kingsville	612	43.0	49	999
Texas A&M San Antonio	524	32.3	42	692
Texas Southern	689	31.8	49	1,074
Texas State	1,538	38.1	37	1,000
Texas Tech	2,096	53.2	53	1,580
UT Arlington	1,585	55.3	79	2,479
UT Austin	3,946	62.2	81	2,708
UT Dallas	2,366	41.9	74	2,014
UT El Paso	1,228	51.2	49	874
UT Permian Basin	455	33.2	72	2,036
UT San Antonio	1,902	54.1	61	1,600
UT Tyler	600	25.7	38	1,230
West Texas A&M	631	43.1	35	1,107
Florida				
Central Florida	3,326	53.6	56	1,346
Florida	3,184	61.6	68	1,997
Florida A&M	667	48.4	34	837
Florida Gulf Coast	1,325	35.2	40	870
Florida International	2,879	46.7	54	1,501
Florida State	3,065	52.4	53	1,518
North Florida	1,262	45.1	36	660
West Florida	1,077	31.7	45	977
Georgia				
Albany State	415	30.6	50	1,212
Augusta	886	34.3	43	1,016
Clayton State	274	21.5	17	225
Georgia	3,378	49.9	62	1,995
Georgia College	483	51.6	29	584
Georgia Southern	1,216	49.9	47	971
Georgia State	1,795	47.6	52	1,494
Georgia Tech	1,549	67.7	121	3,727
Kennesaw State	2,570	39.2	39	805
Middle Georgia State	451	9.5	33	944
North Georgia	839	27.5	21	375
West Georgia	870	22.9	20	232
Indiana				
IU Bloomington	3,354	46.5	64	1,706
Other States				
Arizona State	4,420	46.6	80	2,283
North Carolina Central	722	26.5	24	620
Western Kentucky	1,192	37.0	37	669
Total	**78,383**	**47.2**	**61**	**1,657**

Note: Table B2 presents research productivity for the instructors at each university in the main sample, Fall 2024 to Fall 2026. Of the 78,383 distinct instructors, 36,997 have an accepted research profile. Each instructor teaches at one university in the main sample, so the university rows add up to the total. I compute publication and citation means over matched instructors with a defined value.

provides a binary publication measure and the log h-index for matched researchers. I measure career experience using recovered PhD year, earliest publication year, and, for Texas instructors, hire year, and I include missing-data indicators for covariates with incomplete coverage.

B.3 Course Classification and Text Measures

To compare courses across the 50 universities, I map institution-specific departments and course numbers into common classifications, and I complement them with a set of measures constructed from the syllabus text itself. I describe each in turn.

B.3.1 Harmonization of College Majors

Across the 50 universities in the estimation sample, I observe 7,472 distinct department codes, reflecting different administrative naming conventions across states and institutions. A code such as "BIO" may appear in one place, while another university uses "BIOL" or related prefixes, so a systematic mapping is needed before cross-university comparisons make sense. I harmonize these codes into 42 named categories using the framework of Biasi and Ma (2022), with direct mappings for most codes and university-specific overrides where the same prefix refers to different fields at different institutions. For composite or cross-listed courses, I assign the course to the first-listed department, and this procedure gives me a harmonized major for 98 percent of syllabi. The remaining records fall into a residual "Other" category.

For broader comparisons, I aggregate the 42 harmonized majors into five macro-fields, namely STEM, Humanities, Social Sciences, Business, and Other. The mapping follows mostly straightforward disciplinary boundaries, although a few boundary cases require CIP-based judgment.[1] Table B3 lists the full mapping from department codes to harmonized majors.

B.3.2 Major Characteristics

I complement the syllabus-level data with external major-level covariates that capture labor market conditions, occupational AI exposure, and student selection into each field. These covariates allow me to examine which characteristics of a major predict cross-field variation in AI classroom integration and also serve as controls in specifications that isolate the instructor-level component of AI adoption.

I draw on the Federal Reserve Bank of New York's College Labor Market data (Abel et al., 2014) to recover unemployment rates, underemployment rates, early- and mid-career median wages, and the share of graduates who go on to earn a graduate degree. I also

[1] I classify Education and Criminal Justice as Social Sciences following CIP conventions, Nursing as STEM following NSF and CIP usage, and Interdisciplinary Studies as Humanities based on the composition of its largest constituent departments.

Table B3: Harmonized Majors: Sample Distribution and Illustrative Sub-Fields

		Illustrative Sub-Fields		
Harmonized Major	N	(1)	(2)	(3)
Panel A: Business (N = 31,321)				
Business Administration	15,693	Management	Supply Chain	Entrepreneurship
Accounting	4,713	Auditing	Managerial Accounting	
Finance	4,289	Real Estate	Risk Management	Financial Planning
Marketing	4,076	Sales	Digital Marketing	
Hospitality	2,550	Tourism	Hotel Management	Culinary Arts
Panel B: Humanities (N = 93,077)				
English	15,017	Creative Writing	Literature	Comparative Literature
Music	13,862	Music Education	Music Theory	Music Performance
Communication	13,530	Journalism	Advertising	Public Relations
Art	12,337	Art History	Graphic Design	Digital Media
Interdisciplinary Studies	11,260	Honors	Gender Studies	Global Studies
Foreign Languages	9,388	Spanish	French	Linguistics
History	6,635	American History	European History	
Film & Theater	5,840	Theater	Film Studies	Drama
Philosophy	3,636	Religious Studies	Ethics	
Dance	1,572			
Panel C: STEM (N = 128,831)				
Health Sciences	16,396	Kinesiology	Speech Pathology	Occupational Therapy
Mathematics	15,836	Statistics	Actuarial Science	Applied Mathematics
Engineering (Other)	15,267	Mechanical Eng.	Aerospace Eng.	Industrial Eng.
Biology	14,682	Microbiology	Neuroscience	Biochemistry
Computer Science	11,151	Data Science	Cybersecurity	Software Engineering
Chemistry	7,762	Biochemistry	Medicinal Chemistry	
Nursing	7,092	Graduate Nursing		
Agriculture	5,752	Agronomy	Animal Science	Food Science
Information Systems	5,648	Information Technology	Management Info. Systems	
Public Health	5,518	Pharmacy	Nutrition	Epidemiology
Physics	5,357	Astronomy	Space Science	
Electrical Engineering	5,146	Computer Engineering	Electronics	
Architecture	3,745	Urban Planning	Interior Design	Landscape Architecture
Civil Engineering	3,106	Construction	Water Resources	
Environmental Science	2,552	Sustainability	Oceanography	Marine Science
Geology	2,428	Earth Science	Meteorology	Atmospheric Science
Biomedical Engineering	1,393	Bioengineering		
Panel D: Social Sciences (N = 67,955)				
Education	20,877	Curriculum & Instruction	Special Education	Reading Education
Psychology	11,988	Cognitive Science	Clinical Psychology	Behavioral Analysis
Political Science	6,904	International Relations	Public Policy	Government
Criminal Justice	6,215	Criminology	Forensic Science	Law Enforcement
Social Work	5,903	Counseling	Human Development	Family Studies
Economics	4,087	Agricultural Economics	Real Estate	
Sociology	3,711	Demography	Social Stratification	
Public Administration	3,551	Public Policy	Urban Planning	
Geography	2,397	Geographic Info. Systems	Urban Geography	
Anthropology	2,322	Archaeology	Biological Anthropology	
Panel E: Other (N = 6,155)				
Other	6,155	Legal Studies	Family & Consumer Sciences	

Note: Table B3 presents the distribution of syllabi in the main sample across the 43 harmonized majors, along with illustrative sub-fields that map to each category. Each major appears in the macro-field that covers most of its syllabi. Majors are sorted by macro-field, then by descending count within each field.

measure each major's exposure to language modeling advances using the LMOE index from Felten et al. (2023), assigning occupation-level scores to majors through a CIP-to-SOC crosswalk weighted by BLS employment counts.

From the 2019–20 National Postsecondary Student Aid Study, I recover mean SAT composite scores, mean ACT scores, and mean college GPA for 23 broad major fields. At the institution-by-major level, IPEDS Completions and the College Scorecard provide degree

counts, the share of female graduates, and post-graduation median earnings. In the main heterogeneity specifications, I focus on six covariates, namely LMOE, unemployment, mid-career wage, graduate-degree share, mean SAT, and share female.

B.3.3 Course Levels

I classify each course into one of three academic levels using the first digit of its course number, following Biasi and Ma (2022). Under this convention, courses numbered 100 through 299 are basic undergraduate, those numbered 300 through 499 are advanced undergraduate, and those numbered 500 and above are graduate. This heuristic follows a longstanding convention across public universities and covers the vast majority of the syllabi in the sample.

For the main specifications, I refine this classification into five tiers that distinguish freshman, sophomore, junior, senior, and graduate courses. I construct this finer classification for all 50 universities, with institution-specific adjustments where numbering conventions differ.

For instance, UT Austin uses a three-digit numbering system in which the first digit encodes credit hours rather than academic level. For the Texas universities in the Texas Common Course Numbering System, the four-digit course number also encodes credit hours in its second digit. For the Florida universities, I use the Statewide Course Numbering System, which provides comparable identifiers across institutions.

Across the full sample, 32 percent of syllabi correspond to basic undergraduate courses, 47 percent to advanced undergraduate, and 21 percent to graduate courses. Within the undergraduate tiers, freshman and sophomore courses account for 15 and 16 percent of the sample, while junior and senior courses account for 24 and 23 percent.

B.3.4 Course Characteristics

For four universities in the sample, I recover section-level grade distributions from publicly available registrar records. I use UT Austin and UT Dallas grade repositories, Texas A&M registrar files, and Houston's public grade records. From these records, I compute mean GPA on a 4.0 scale, the share of students receiving a D, F, or Withdrawal, and the share receiving an A.

I also recover section-level enrollment or capacity for 44 of the 50 universities from registrar and schedule databases, including IU Bloomington, Houston, Florida International, and Georgia. Because these sources differ somewhat in coverage and precision across universities, I combine them into a harmonized enrollment measure that uses observed enrollment when available and otherwise relies on seat counts, current capacity, or historical averages.

I also recover semester credit hours for most of the sample. At the Texas universities, credit hours are encoded in course numbering conventions, and elsewhere I draw on registrar schedules and the syllabus text. Credit hours are available for 87 percent of the syllabi

in the sample, and I include a missing indicator for the remaining syllabi in every specification that uses this variable.

B.3.5 Syllabus Text Characteristics

I construct nine text-based controls from the syllabus text, capturing the complexity of the document's prose, the prescriptiveness of the instructor's tone, and the presence of explicit course policies. The first three capture text complexity, namely the Flesch-Kincaid Grade Level, the Moving-Average Type-Token Ratio, and average sentence length. I additionally include the log of total word count, which absorbs the mechanical correlation between syllabus length and the likelihood of detecting any keyword-based outcome. Across the 50 universities, average syllabus length ranges from 1,206 words at North Carolina Central to 5,480 words at Texas A&M International.

I also capture instructor tone through the prescriptive modal share, penalty language density, supportive language density, and binary indicators for explicit late-work and attendance policies. All continuous text controls are standardized to mean zero and unit standard deviation in the estimation sample. I recover all nine measures for all but 109 syllabi in the main sample.

I also identify whether a teaching assistant is mentioned in each syllabus, combining text-based detection with structured platform metadata where available. This approach identifies a TA in 14 percent of syllabi. Table B4 shows definitions and summary statistics for the nine text controls.

Table B4: Text Control Variables: Summary Statistics

Variable	Description	Mean	SD
Flesch–Kincaid Grade	Readability index (higher = more complex)	14.56	2.45
Lexical Diversity (MATTR)	Moving-average type–token ratio (50-word window)	0.72	0.03
Avg. Sentence Length	Mean words per sentence	21.72	4.40
Modal Prescriptive Share	Share of modal verbs that are prescriptive	0.28	0.13
Penalty Language Density	Penalty-related terms per 1,000 words	2.32	2.08
Supportive Language Density	Supportive/encouraging terms per 1,000 words	2.88	2.09
Has Late-Work Policy	Binary indicator for a late-work policy	0.62	0.49
Has Attendance Policy	Binary indicator for an attendance policy	0.69	0.46
Log Word Count	Natural log of full-text word count, top-coded at 9,500	7.84	0.63

Note: Table B4 presents summary statistics for the nine text-derived control variables. I compute all metrics directly from the syllabus full text. Flesch–Kincaid Grade and Avg. Sentence Length are winsorized at the 99th percentile to limit the influence of syllabi lacking sentence punctuation.

B.3.6 AI-Related Content

To identify syllabi that contain AI-related content, I use ten keyword patterns in two tiers. The first covers policy-oriented post-ChatGPT terms, and the second covers broader references such as "artificial intelligence," "AI," "generative AI," and "large language model." I exclude terms with high false-positive rates and those that add little signal beyond these core patterns. The standalone acronym "AI" requires special treatment because it often

refers to "Associate Instructor," building codes, and room numbers, so I apply a contextual filter blocking known false positives and requiring a nearby AI-related term. I use the resulting indicator of any AI mention to compute the AI-mention rates I present in Section 2 of the main paper. The four classification outcomes are constructed separately from the full syllabus text, as I describe in Appendix C.

Appendix C: Classification Pipeline Details

This appendix provides the complete coding guidance for each of the four classification measures, including all worked examples of affirmative and negative cases drawn from actual syllabus text (Tables C1 and C2). The classification models and the human raters in the validation exercise all received this same guidance.

Table C1: Policy and Adoption Questions: Full Classification Guidance

Has AI Policy

Question. Does the syllabus contain an explicit policy, rule, or guideline about student use of AI or generative AI tools (such as ChatGPT, Copilot, or similar tools) for coursework?

Yes if the syllabus states rules, permissions, prohibitions, guidelines, or expectations about whether and how students may use AI tools for coursework. The policy may come from the instructor or the university and may appear anywhere in the syllabus. It may permit, prohibit, restrict, or conditionally allow AI use.

(i) "AI tools such as ChatGPT may not be used in this course"
(ii) "Students are welcome to use AI tools for assignments"
(iii) "Per university policy, AI-generated content must be cited"
(iv) "Use of generative AI is prohibited unless otherwise stated"
(v) "You may use AI for brainstorming but not for final submissions"
(vi) "All work must be your own; AI-generated submissions will be treated as plagiarism"

No if the syllabus teaches AI as a course topic but says nothing about whether students may use AI tools for their own coursework, or contains only a generic academic integrity statement that does not specifically mention AI or generative AI tools, or has no AI mention.

Permits AI

Question. Does the syllabus explicitly state that students are permitted, allowed, or welcome to use AI tools for coursework in this course?

Yes if the syllabus contains an explicit statement permitting, allowing, welcoming, or approving student use of AI tools for coursework. The permission may appear in a policy section, assignment description, course overview, or anywhere in the syllabus. It can be broad or targeted to specific assignments.

(i) "Students are welcome to use AI tools in this course"
(ii) "AI tools such as ChatGPT may be used for assignments"
(iii) "You are free to use generative AI"
(iv) "Per university policy, AI use is permitted with proper citation"
(v) "AI is allowed for brainstorming and drafting"

No if the syllabus only prohibits or bans AI without any permission, discusses AI only as a topic to study without permitting student use, requires AI disclosure without explicitly saying AI is permitted, or has no AI mention.

Note: Table C1 presents the complete coding guidance for the two policy questions. I classify these questions from the full syllabus text.

Prompt Structure and Model Configuration. I submit each of the four classification questions independently, so that a model's response to one question does not shape its evaluation of the remaining questions (Haaland et al., 2025; Saltiel, 2026). For each question, I provide the model with the full text of the syllabus and instruct it to base its response only on what is explicitly stated in the document.

For each syllabus-question pair, I construct a user message that inserts the question text, the corresponding coding guidance from Tables C1 and C2, and the full syllabus text. In

Table C2: Integration Questions: Full Classification Guidance

Teaching Integration

Question. Is AI part of what this course teaches: its class sessions, topics, or course materials?

Yes if the course dedicates class time or course content to AI: a schedule, module, or session entry on an AI topic, a class activity, lecture, discussion, or demonstration involving AI, or an assigned reading, video, tutorial, or other course material about AI. One concrete line is enough: a single schedule or module entry naming an AI topic or activity counts.

(i) "Week 7: AI in Public Administration" (a schedule entry)
(ii) "AI Ethics Discussion" (a one-line module element)
(iii) "In class, we will use ChatGPT to critique a sample essay" (a class activity)
(iv) "Watch: How Large Language Models Work" (an assigned reading or video about AI)

No if AI appears only outside the taught content: only in graded-task requirements with no class time or course material on AI (that is assessment, not teaching), only in rules about student use, only in a learning objective or general statement that the course will involve AI, or "AI" that does not mean artificial intelligence. Statements about AI use inside the course's AI policy or rules section are policy, not course content, even when they mention class exercises.

(i) "AI Analysis Project (20%): use ChatGPT to draft a memo" (graded work only, no AI class content)
(ii) "You may use AI tools if you cite them" (a use rule)
(iii) "Generative AI will be integrated throughout this course" (a general statement)
(iv) "Appreciative Inquiry (AI) reflection due Friday" (not artificial intelligence)

Evaluation Integration

Question. Does at least one graded or assessed task require students to engage with AI?

Yes if a specific assignment, project, exam, or task requires AI in the student's own work: the task requires students to use an AI tool, the deliverable must include or build on AI output, or a required step has students produce, compare, critique, or revise AI content. Look for the requirement in the task itself. A one-line graded item whose title names AI counts, unless the task is clearly about AI as a topic rather than using it. Course objectives are not tasks.

(i) "Homework will require the use of AI tools for problem solving" (a task requiring AI use)
(ii) "AI Analysis Project (20%): use ChatGPT to draft a memo" (a named graded assignment)
(iii) "Assignment 4: submit the transcript of your ChatGPT conversation" (AI output in a graded deliverable)
(iv) "Graded practice (15%): motivational interviewing with ChatGPT" (a required interaction)
(v) "Assignment 2 (10%): generate a response with ChatGPT and critique its accuracy" (a required critique step)
(vi) "Final project: build and train a machine learning model" (a build requirement)
(vii) "AI Assignment (10%)" (a one-line graded item)

No if no assessed task requires AI: optional, invited, or extra-credit use, course-wide permission, citation or disclosure rules, and task-like statements inside the AI policy section. Graded work that only tests AI as subject matter, or "AI" not meaning artificial intelligence, is also No.

(i) "You may use AI for the research paper" (optional use, even on a specific task)
(ii) "If you use AI, include your prompts and its responses" (a disclosure rule)
(iii) "The midterm covers machine learning concepts" (AI as subject matter)

Note: Table C2 presents the complete coding guidance for the Teaching Integration and Evaluation Integration questions. I classify both questions from the full syllabus text with AI policy statements removed.

addition to a binary classification, I require each model to provide a confidence rating, a brief justification, and a direct quotation from the syllabus when the answer is affirmative. This allows me to assess both the classification and the quality of the model's reasoning. Table C3 presents the user message template verbatim.

Table C3: User Message Template for Syllabus Classification

USER MESSAGE TEMPLATE

QUESTION: {question text}

CODING GUIDANCE: {guidance text}

SYLLABUS TEXT: {full_text}

INSTRUCTIONS
1. Read the FULL syllabus text carefully.
2. Answer the question with Yes or No.
3. State your confidence: high, medium, or low.
4. Provide a 2–3 sentence justification for your answer.
5. If your answer is Yes, quote the MOST RELEVANT passage(s) from the syllabus as evidence. If No, write "N/A".

Return ONLY a single JSON object with exactly these keys:

```
{
  "answer": "yes" or "no",
  "confidence": "high" or "medium" or "low",
  "justification": "your explanation here",
  "evidence": "quoted text or N/A"
}
```

Note: Table C3 presents the user message template populated for each syllabus-question pair. Placeholders in braces are replaced with the specific question text, coding guidance, and full syllabus text.

I classify the two integration questions with three open-weight models run locally, Gemma-4-31B, Qwen3.8-27B, and Muse-Glimmer-30B, and take the label to be the modal answer of the three. For each of the three models, I set the temperature to zero and fix the random seed, aiming to maximize deterministic outputs across repeated runs.[1] Before classifying the two integration questions, I identify the AI policy statements in each syllabus and remove them from the text. The integration questions therefore read the course content itself, its schedule, activities, readings, and assignments, rather than the policy language about whether students may use AI. The two policy questions, where this language is itself the outcome of interest, are classified on the full syllabus text.

Human Validation. I recruited four research assistants to independently classify a sample of 60 syllabi across the full set of four measures.[2] I drew the syllabi using stratified sampling

[1] I also set the presence penalty and frequency penalty to zero and top_p to one for all three models.
[2] The rating design assigns 20 core syllabi to all four raters and rotates the remaining 40 so that each rater

by AI mention status, institution, and classification profile to ensure representation across the distribution of instructor responses. Each rater received the complete coding guidance (Tables C1 and C2), a brief orientation document describing the classification task, and the full text of each syllabus in PDF format. The raters, working independently, agree closely both with one another and with the modal classification, which comfortably passes the Alternative Annotator Test.

Cross-Model Agreement. I classify the two integration questions with three open-weight models and take the label to be the modal answer of the three. Table C4 reports the pairwise agreement rates for the two questions. The three models agree on the large majority of syllabi, with agreement beyond chance highest for teaching integration and somewhat lower for evaluation integration.

Table C4: Cross-Model Agreement on the Integration Questions

	Gemma-4-31B vs. Qwen3.8-27B		**Gemma-4-31B vs. Muse-Glimmer-30B**		**Qwen3.8-27B vs. Muse-Glimmer-30B**	
	Agreement (%) (1)	Cohen's κ (2)	Agreement (%) (3)	Cohen's κ (4)	Agreement (%) (5)	Cohen's κ (6)
Teaching Integration	98.1	0.798	98.2	0.826	98.6	0.835
Evaluation Integration	98.5	0.687	98.3	0.660	99.3	0.803
Syllabi, Teaching Integration			325,756			
Syllabi, Evaluation Integration			326,541			

Note: Table C4 presents, for each pair of classifiers, the share of syllabi on which the two give the same answer, in percent, and Cohen's κ. Syllabi that never mention AI, or mention it only in the AI policy, are coded as not having each measure for every classifier and are included in the comparisons. Each pair is computed on the syllabi rated by both classifiers, and the Syllabi rows give the smallest of the three pair counts. AI Policy and AI Permitted come from a single classifier and have no pair.

Qualitative Categories. The classification described above establishes whether a course integrates AI into its teaching or its evaluation, but it is silent on what that integration actually involves. A course that devotes a module to neural networks and a course that asks students to draft essays with ChatGPT both count as integrating AI, yet they reflect very different classroom choices. To distinguish between them, I build a qualitative layer on top of the classification. The natural input for this exercise is the passage of syllabus text that the classifier quotes as evidence when it answers a question affirmatively, since working with the quoted passage isolates the exact language describing how AI enters the course, rather than the syllabus as a whole.

Given these passages, the sorting itself is deliberately simple. I define one fixed list of words and phrases per category, and a passage joins a category when it contains at least one term from that category's list. The list for AI Methods, for instance, includes machine learning, neural networks, and deep learning, while the list for Using AI Tools includes

classifies 50 syllabi in total.

use AI and AI tools. A passage can join several categories, and passages that match no list remain in a residual group. As such, no model judgment enters this step.

On the teaching side, the categories describe what the AI content of a course is about, distinguishing courses that teach the technology itself from courses in which students work with AI tools and courses that apply AI to the course's own subject. On the evaluation side, they instead describe what students are asked to do with AI in graded work, whether building AI models, engaging with AI output, or producing their own assignments with AI assistance. In this way, the categories translate a binary signal of integration into a description of the choices instructors are actually making.

The lists must sort passages the way a careful reader would, and I verify this in a blind audit of teaching passages, checking that the category assigned by the lists matches the one a reader would choose after reading each passage in full. Moreover, the terms most distinctive of each category's passages line up closely with what each category is meant to capture. Altogether, these exercises establish that the passages are sorted sensibly and that the categories are distinct in their language.