How exposed is higher education to Artificial Intelligence?

Jacob Light*

March 26, 2026

Preliminary - please do not cite or circulate
[Click here for latest version]

Abstract

How exposed is higher education to artificial intelligence, and how are institutions and students responding? I construct a novel measure of curricular exposure to AI by applying the task-based framework of Webb (2019) to a comprehensive dataset of course offerings from more than 1,000 U.S. colleges and universities. Extracting verb-object pairs from course descriptions and patent texts, I measure the overlap between what students are trained to do and what AI and other technologies can do. I find that college curricula are more exposed to AI than to prior technological shocks, and that exposure has increased over the past decade, driven by the diffusion of data analysis coursework across disciplines. Exposure is highest in fields such as Statistics, Economics, and Computer Science, and at research universities. Despite this high and growing exposure, I find limited evidence of adjustment: using a panel of course syllabi spanning the pre- and post-ChatGPT periods, instructors rapidly adopt AI policies but make only modest changes to course structure, and student enrollment patterns show little response. These findings suggest that higher education may adjust slowly to technological change, even when the skills it produces are highly exposed.

JEL Classification: J24, I2, O3.

*Hoover Institution, Stanford University. jdlight@stanford.edu

1 Introduction

Technological disruption has long been a feature of labor markets. Each wave of innovation expands the production possibilities frontier but may simultaneously render specific skills — and the workers who supply them — less valuable. In recent decades, the brunt of automation has fallen on routine, lower-skill jobs (Autor et al. 2003). Under such skill-biased technological change, a standard policy prescription has been to encourage workers at risk of displacement to upskill through higher education or lifelong learning on the premise that higher education equips workers with harder-to-automate capabilities. That logic may now be in question.

A new generation of artificial intelligence (AI) tools differs from earlier technologies in three notable ways. First, their potential application space is extraordinarily broad, spanning language, vision, prediction, and code. Second, many models operate with minimal human supervision, performing entire task bundles rather than single modular steps. Third — and central to this paper — the tasks at which modern AI excels overlap heavily with those performed by workers in traditionally high-skill occupations (Brynjolfsson and Mitchell 2017). If AI increasingly competes on the same cognitive terrain that universities teach, the return to "upskilling" through college coursework may erode faster than in previous eras.

This paper provides the first systematic measurement of how exposed college coursework is to modern AI capabilities. Building on the task-based framework of Webb (2019), I combine a novel dataset of course descriptions, enrollments, and offerings from 1,019 U.S. colleges and universities (Light 2026) with two measures of technological capabilities. The first uses patent data to characterize the distribution of AI capabilities at the technological frontier. The second, and preferred measure, uses task-level estimates of exposure to large language models (LLMs) from Eloundou et al. (2023), capturing the capabilities of tools that are already widely used by students and instructors.

For each field of study, I extract verb-object pairs (e.g., "analyze data," "write essays") from course descriptions, construct a task distribution, and measure its overlap with the task distributions implied by each technology. The resulting *exposure score* captures the extent to which tasks taught in college can be performed more efficiently using a given technology. The measure is intentionally agnostic about whether AI substitutes for or complements these tasks.

The paper documents three main findings. First, college coursework is substantially more exposed to AI than to earlier waves of technological change. While exposure to robotics remains low across fields, exposure to AI — particularly under the LLM-based measure — is widespread, with many fields lying well above the corresponding distribution of occupa-

tional exposure. Second, exposure to AI has risen steadily over the past decade, driven by the diffusion of data analysis and related tasks across disciplines. Third, exposure varies meaningfully across institutions within the same field, reflecting differences in course offerings and curricular depth rather than differences in how individual courses are taught.

In the second part of the paper, I examine how higher education responds to a sharp technology shock: the release of ChatGPT in November 2022. The availability of LLM tools lowers the cost of completing coursework in tasks that overlap with AI capabilities but also introduces uncertainty about the labor market returns to those same skills. I study responses along both the supply and demand margins.

On the instructor side, I construct a panel of syllabi covering all courses offered at more than 20 colleges and universities to track changes in course design. I find rapid adoption of AI policies — by Fall 2025, a majority of courses explicitly regulate the use of AI tools, with higher rates of restriction in the Humanities and Social Sciences than in STEM or Business — but relatively modest changes in course structure. In particular, there is only limited reweighting toward in-class assessments such as exams and participation. These adjustments are small and gradual, especially when compared to the large and persistent changes in course design observed during the COVID-19 pandemic. On the student side, I measure enrollment responses across fields with different levels of AI exposure. Despite the rapid diffusion of LLM tools, I find little evidence of reallocation in course-taking behavior in the short run. Enrollment patterns remain largely continuous through the post-ChatGPT period.

Taken together, these findings point to a gap between exposure and response. While college coursework is highly exposed to AI broadly, and LLM capabilities specifically, both instructors and students exhibit relatively limited short-run adjustment. Contrasts between the pandemic — where instructors responded in dramatic and persistent ways to an instructional environment where monitoring student learning through exams became more difficult — and the ChatGPT period — where new difficulties emerge in evaluating student learning through out-of-class assessments, but no comparable changes emerge — are striking. These contrasts suggests that institutional frictions, uncertainty about the long-run role of AI, and the difficulty of redesigning curricula may slow the responsiveness of higher education to technological change.

I validate the task-based measurement approach using several complementary exercises. The tasks extracted from course descriptions align closely with intuitive differences across fields. For example, "analyze data" and "solve equations" are concentrated in quantitative disciplines, while "write essays" and "interpret texts" are more prominent in the Humanities. Moreover, the task distributions implied by course descriptions correlate strongly with

the tasks performed in the occupations associated with each field. These patterns suggest that the exposure measure captures meaningful variation in the content of human capital formation.

Taken together, the results document a striking pattern: while AI capabilities overlap substantially with the content of higher education, institutional responses along both the extensive and intensive margins remain modest in the short run.

The paper makes three central contributions. First, it extends task-based exposure metrics from occupations to the human capital formation in college, offering a forward-looking benchmark for policy discussions around the value of college majors during a period of potential AI disruption in the labor market. Second, the analysis reveals substantial within-major heterogeneity across institution types, driven largely by differences in the level of courses offered rather than differences in skills/tasks emphasized in the same course offered across institutions. The findings demonstrate how the "same degree" can differ meaningfully depending on where it is earned and reveal limitations of data sources that record information only on an individual's field of study. Third, it provides new evidence on how both course supply and student demand respond to a fast-moving technological shock.

The paper builds on a large literature in labor economics studying how technology affects employment and occupational choice. Autor et al. (2003) introduced the idea of modeling occupations as bundles of tasks, with technology potentially complementing or substituting for workers in those tasks. Subsequent work has shown that different technologies vary in how they alter demand for skill — some technologies complement cognitive or non-routine work, while others automate routine tasks. For example, technological change in the late 20th century was skill-biased: it disproportionately substituted for workers in jobs with lower education requirements while complementing highly educated workers, thereby contributing to the growing earnings gap between more and less educated workers (Autor et al. 2006). Although evidence on the labor market effects of AI remains nascent because of its recent development and, as yet, limited deployment (McElheran et al. 2024), there is growing empirical support for these task-based predictions. Establishments that have adopted AI tend to demand more workers who can integrate AI into production processes, while simultaneously reducing demand for workers in occupations that AI can readily perform (Acemoglu et al. 2022).

Several recent papers have developed methods to measure the exposure of occupations to new technologies. Felten et al. (2021) and Brynjolfsson et al. (2018) rely on crowdsourced scoring: Brynjolfsson et al. (2018), for example, use a standardized rubric to assess whether tasks are suitable to machine learning based on how clearly inputs map to outputs and how easily performance can be measured. Felten et al. (2021) focus on perceptions of overlap

between specific AI capabilities and occupational abilities. In contrast, Webb (2019) uses a data-driven approach: he measures exposure by comparing the text of patents (which describe what technologies do) with O*NET task descriptions (which describe what workers do). This method has several advantages for my project. First, it relies on textual similarity between technology capabilities and tasks, rather than subjective judgments of suitability. Second, because it is based on patents rather than realized wage changes or job postings, it provides a forward-looking measure that reflects the technological frontier, rather than current market conditions. Finally, the method can be easily updated as new patents emerge, making it a flexible way to track technological change over time.[1]

My paper applies this logic one step earlier in the labor market pipeline — mapping patents not to occupations, but to the tasks embedded in college courses. This shifts the unit of analysis from workers to students and moves focus from firms' hiring decisions to universities' curricular decisions and students' human capital investment.

Applications of a task or skills framework to research in higher education are more limited, largely because of the need for rich data that records skills or tasks at a field- or occupation-by-field level. Recent work in this area has relied on either course descriptions or course syllabi to extract institution- and course-level detail on course content. Closest in spirit to my project is work by Javadian Sabet et al. (2024), who link course descriptions from a large corpus of syllabi at numerous colleges and universities to detailed work activities in the labor market. Their data tool offers an alternative source for assessing course or field exposure to AI, although it is limited for measuring exposure of more recent courses to technological change because its data cutoff is 2017. Biasi and Ma (2022) also use this syllabus data to track how courses incorporate material at the frontier of academic research. While this shares my goal of linking course content to cutting-edge knowledge, their focus is on courses' inclusion of academic frontier material, whereas my analysis asks whether courses incorporate tasks that overlap with what technologies at the frontier can do. Finally, a recent white paper by Timmerman (2025) measures field-level AI exposure by mapping majors to occupations and applying occupational AI exposure scores. My approach differs in measuring exposure directly at the field-of-study level, rather than using occupations as a mediator. It also benefits from rich data that allow me to measure heterogeneity in exposure within the same field across institutions and over time.

The rest of the paper proceeds as follows. In Section 2, I summarize the data sources used for this project. Section 3 outlines the text analysis used to measure field-technology

[1] Although I follow Webb's use of patents here, the same methodology could also be applied to other sources of technology data, like GitHub commits or academic papers, to expand the measure's coverage as AI evolves.

exposure. I walk through validation exercises in Section 4 to stress-test the task-based framework as applied in this new setting. Section 5 summarizes results in the measurement of field-technology exposure, and Section 6 summarizes instructor and student responses to ChatGPT's release. Section 7 concludes.

2 Data

The primary data sources for this project are patent text from the US Patent Office, covering all patents filed between 2000-2025, and comprehensive course offerings data from 1,019 US colleges and universities.

2.1 Course data

2.1.1 Course catalog data

The course catalog dataset contains detailed information on course offerings at a large, nationally representative panel of US colleges and universities (Light 2026). I collected these data by scraping online course catalogs and course schedules, capturing 50 million course sections offered since 1996 across 1,019 institutions.

The dataset forms an unbalanced panel that includes the complete set of course offerings for each institution-year. For each course section, I observe enrollment, instructor(s), instruction format (e.g., in-person or virtual), and a brief text description of course content. Figure 1 is an example of the information contained in a typical observation. Most crucial for this project are the course descriptions, from which I extract tasks, and the enrollment counts.

The sample is broadly representative of US colleges and universities. It covers 37% of non-profit four-year institutions and 31% of non-profit two-year institutions, including 58% and 43% of enrollment, respectively. It also mirrors national distributions of key characteristics — such as selectivity, tuition costs, and endowment levels — although it skews somewhat toward larger, public institutions because it misses many very small, often religiously affiliated private colleges.[2]

2.1.2 Syllabus data

Measuring the evolution of coursework in response to ChatGPT requires detailed information on the structure of college courses. I construct a panel of course syllabi from 20 US colleges and universities. Institutions in the syllabi sample cover a range of institutions, ranging

[2]See Light (2026) for a detailed discussion of coverage and representativeness.

from highly selective to non-selective four-year institutions, covering data pre- and post-the release of ChatGPT (and as early as 2005). The sample contains publicly available syllabi from universities that publish full syllabus archives online, and not behind a login wall. Essentially, I only include institutions for which as complete a set of syllabi as possible are available. The majority of the schools in the syllabus sample are located in Texas; Texas HB 2504 requires universities to make publicly available course syllabi and other course details. While the law requires archives of only two years, many institutions have archives that date back to the law's passage in 2009.

For this paper, I use an automated approach to extracting from the syllabus related to course structure and administration. In particular, I extract course identification information (course name, ID, title, and period of course offering), instructor information, grade components, and various course policies using the OpenAI API.

For the data to be useful, it is essential that the data extraction match the information in the syllabus. Of particular concern is the risk of AI hallucinations producing output that seems plausible but is not accurate. I take a few steps to guard against this. First, I have the LLM output raw text from the syllabus in addition to the processed grade component name and weight. I can, therefore, confirm that the text output matches exactly what appears in the syllabus. Second, I validate the accuracy of the LLM output using a human evaluator based on a sample of 250 syllabi. The LLM accurately extracts the full grading scheme in 85% of cases.[3]

I classify the grading components into seven categories using string matches on the raw text of the grading component: exams (tests, quizzes, etc.), homework, participation, projects, presentations, labs, and essays. I classify components that do not fit into one of the seven categories into an "other" category.

2.2 Technology data

2.2.1 Patent data

I obtain patent titles and abstracts for all patents filed with the US Patent and Trademark Office between 2000-2024. Titles typically provide brief descriptions of a technology's core function, while abstracts provide longer summaries of the technology's development and applications. For my analysis, I use patent titles to extract the tasks a technology can

[3]Errors typically take the form of adding an extra grading component that is the total of different subcomponents: for example, a syllabus may list Assignment 1 and Assignment 2 as grade components and include an Assignment Subtotal line in the grade composition table. In post-processing, I remove any grade item that includes the string "total." These errors do not appear to be systematic for specific grading components or in specific fields or years. Thus, the consequence of these inaccuracies will primarily be to introduce measurement error into the estimates.

perform, following Webb (2019).[4] An example of a patent title is "Analyzing flight data using predictive models" (Desell et al. (2019)). The final patent dataset includes approximately 135,000 patents covering Robotics and Artificial Intelligence.[5]

To classify patents into technology classes, I apply a keyword filter similar to Webb (2019). For Robots, I include patents whose titles or abstracts contain the word "robot" or derivatives of "manipulate," excluding patents with CPC codes A61 (medical or veterinary science) and B01 (physical or chemical processes). For AI, I expand Webb's definition (which pre-dates the release of ChatGPT) to include phrases associated with recent AI innovations: "artificial intelligence," "deep learning," "large language model," "llm," "neural network," "predictive model," "reinforcement learning," "supervised learning," or "transformer model."

The resulting dataset includes 28,165 Robot patents and 26,136 AI patents. Approximately half of the patent titles contain at least one verb-object pair, resulting in around 12,000, 16,000, and 43,000 verb-object pairs for the three technology classes, respectively.

2.2.2 LLM capability data

My second measure of AI capabilities is based on Eloundou et al. (2023), who evaluate the extent to which large language models (LLMs), specifically GPT-4, can reduce the time required to complete tasks in O*NET. Using a combination of crowdsourced human assessments and model-based evaluations, they classify each task into three categories. Tasks are assigned *no exposure* if GPT-4 produces little or no reduction in completion time; *direct exposure* if GPT-4 alone can reduce completion time by at least 50 percent; and *indirect exposure* if achieving this reduction requires GPT-4 in combination with additional tools or software.

To integrate these data into my framework, I follow Webb (2019) and extract verb-object pairs from each O*NET task description. I then map the categorical exposure classifications into a continuous measure by assigning each task a score between 0 and 1, following Eloundou et al. (2023). Finally, I aggregate exposure to the verb-object pair level by averaging across all tasks containing a given pair.

This procedure produces a measure of exposure that captures the extent to which indi-

[4] Although abstracts offer richer detail, they often include background or development information unrelated to direct applications. Results using abstracts are qualitatively similar to those using titles.

[5] I do not report results for software exposure, which is included in Webb (2019). In my data, the distribution of tasks associated with software patents overlaps substantially with those associated with AI patents, particularly in more recent years. This likely reflects both classification challenges and the increasing integration of AI functionality into software-related innovations. As a result, distinguishing software from AI exposure using patent classifications introduces substantial ambiguity. Because the focus of this paper is on exposure to modern AI capabilities — especially those captured by LLM-based measures — I restrict attention to Robots and AI, which remain more clearly separable in the patent data.

vidual task components can be performed more efficiently using LLMs. Unlike patent-based measures, which reflect the distribution of capabilities embodied in technological innovations, this measure directly reflects the ability of currently available models to substitute for or augment specific task inputs.

2.3 Supplemental data sources

For validation exercises, I use occupational task descriptions from O*NET, a database maintained by the US Department of Labor that summarizes worker attributes and job characteristics for nearly 1,000 occupations. These occupation details are gathered through firm and worker surveys to maintain up-to-date information. For my analysis, I use task summaries linked to occupations in the O*NET data.[6] I use O*NET ratings of task importance and relevance to weight verb-object pairs in constructing exposure measures.

To link fields of study to occupations, I use field-occupation mappings estimated from the American Community Survey (ACS). Specifically, I use ACS surveys from 2009-2019 and restrict to employed workers ages 25-65 who hold a four-year degree. This mapping serves as an external validity check for my task-based exposure measures, allowing me to compare the skills emphasized in college courses to those performed in jobs linked to these courses.

3 Methodology

The core methodology for this project involves three main steps: first, extracting "tasks" (verb-object pairs) from course descriptions and patent titles; second, constructing distributions of these tasks for each field of study and technology class; and third, calculating an overlap statistic that reflects the exposure of each field (or field-institution cell) to each technology. This exposure score captures the degree to which the core tasks students perform in their courses are also performed by new technologies.

Throughout, exposure is defined at the verb-object pair level and then aggregated to the course and field level; all reported results reflect these aggregated measures.

3.1 Defining and extracting tasks

Text data encode rich information, but the core challenge is reducing this high-dimensional, interconnected content into a tractable form that captures what is economically meaningful. In this project, I must extract from course descriptions — documents that include administrative, pedagogical, and contextual details — only the parts relevant for measuring a

[6] For example, an O*NET task for Economists is "Compile, analyze, and report data to explain economic phenomena and forecast market trends, applying mathematical models and statistical techniques."

course's technology overlap. A similar process occurs, implicitly, in most economic analyses: for example, researchers often control for an individual's race, gender, or age in labor studies, but not for their quarter of birth or year of marriage unless relevant. With text data, the separation between relevant and irrelevant information is harder to anticipate, so applications typically rely on data-driven approaches to isolate the pieces of text most pertinent for analysis.

I define a *task* as a verb-object pair. For example, consider the description for Stanford's COMM 274D (Public Affairs Data Journalism II): "Learn how to find, create and analyze data to tell news stories with public service impact. Uses relational databases, advanced queries, basic statistics, and mapping to analyze data for storytelling..." (Stanford 2025). The description includes multiple verb-object pairs: "find data," "create data," "analyze data," "tell stories," "use databases," "use queries," "use statistics," and "use mapping." These pairs provide a concise representation of the main activities taught in the course. The motivation for the analysis in this paper is that if patents for new technologies also perform these tasks, the course content itself (and, therefore, the human capital students develop in the course) may be exposed to potential automation or augmentation by that technology.

Given the large size of the text corpus and the complexity of natural language, manual review is not feasible. Instead, I use common Natural Language Processing (NLP) techniques to clean and parse the text at scale. I pre-process the text by removing punctuation and standardizing capitalization. Next, I extract verb-object pairs using the `spaCy` package in Python, a widely used NLP tool (Honnibal et al. 2020). I then remove pairs in which the verb or object is a stopword[7], and lemmatize the verb and object to a common form (e.g., "analyzing" and "analyze" become "analyze").

Course descriptions are noisier than patent titles or O*NET tasks, as they often include administrative details or generic phrases that are not relevant for measuring the human capital students develop in these courses. To reduce noise, I first remove "boilerplate" sentences (such as prerequisites and enrollment restrictions). Next, I calculate chi-squared statistics for each verb-object pair, comparing their occurrence frequency in course descriptions versus patents. I drop pairs that are among the 100 pairs with the highest chi-squared statistic, suggesting that these pairs are broadly related to course descriptions as a document type rather than the specific tasks students perform or human capital they develop.[8]

To further ensure that the verb-object pairs are closely related to course tasks, I project the cleaned task dictionary onto the union of task pairs that appear in either the patent data or the O*NET task inventory (excluding previously dropped pairs). This step ensures that

[7]For example, "use" or "have," which lack meaningful economic content (Jurafsky and Martin 2014).
[8]For example, "introduce student" or "receive credit."

exposure scores only reflect tasks that are relevant to technological change or to job tasks more broadly.

The final dictionary includes 27,612 unique verb-object pairs.

3.2 Exposure measure

I calculate the exposure of a field of study s to a technology class t as the frequency-weighted average of the tasks students perform in that field, weighted by the importance of those tasks to the technology. Let D denote the dictionary of verb-object pairs d. For field s, the share of task d is:

$$x_s(d) = \frac{\sum_{\delta \in D_s} \mathbb{I}(\delta = d)}{|D_s|} \tag{1}$$

where D_s is the set of all task observations in field s. For example, in Economics courses offered in 2023-24, "analyze data" accounts for 1.8% of all verb-object pairs.

Similarly, for each technology class t, I compute the share $w_t(d)$ of task d among all tasks mentioned in patents for that technology:

$$w_t(d) = \frac{\sum_{\delta \in D_t} \mathbb{I}(\delta = d)}{|D_t|} \tag{2}$$

The exposure of field s to technology t is then the weighted average of task shares:

$$exposure_{s,t} = \sum_{d \in D} x_s(d) w_t(d) \tag{3}$$

Intuitively, this measure captures how much of the "task footprint" of a given field overlaps with the tasks that a given technology is already able to perform. Higher values of $exposure_{s,t}$ indicate that a larger share of the tasks emphasized in a field's coursework are tasks that new technologies claim to be able to do. In later analyses, I extend this exposure calculation to field-institution cells to account for differences in course sequencing and teaching emphasis across different types of colleges and universities.

4 Validation

I conduct a series of validation exercises to confirm that the task-based exposure measures capture meaningful differences across fields of study and align with intuitive and observable patterns in the data.

4.1 Distinctive tasks by field of study

To demonstrate how the task extraction approach captures variation across fields, Table 2 summarizes the most common verb-object pairs for a representative sample of fields. For each field, I select the three verbs that appear most frequently and, for each verb, list the (up to) four most frequent objects.

The table highlights distinctive and intuitive task patterns across fields. For example, English courses emphasize tasks such as "reading texts" and "writing essays," Education courses emphasize "teaching lessons" and "developing curriculum," and Arts courses emphasize "creating images" and "performing music." At the same time, some tasks — particularly those involving data analysis — appear across multiple fields, reflecting the growing role of data skills across a range of disciplines.

Columns 5-7 of Table 2 show the technology exposure of these most important tasks, calculated as a weighted average of task frequency in course descriptions and task relevance to each technology class. For readability, I multiply the scores by 100. The table reveals that data analysis tasks are highly exposed to AI across the board, with particularly high exposure for fields such as Economics and Statistics/Data Science where this form of analysis is emphasized. More broadly, exposure to Robots is generally the lowest across all fields. Where tasks overlap with AI, the extent of that overlap is often much greater than for other technology classes.

4.2 Linking field tasks to occupation tasks

As a second validation exercise, I assess whether the tasks extracted from course descriptions for a given field also appear in the occupations into which graduates from that field typically flow. This provides an external check that the task-based measure of field exposure captures tasks that matter in the labor market.

To do so, I use field-to-occupation mappings developed from the ACS for prime-age workers (ages 22-65).[9] I then construct task vectors for each occupation and each field of study, based respectively on the importance of verb-object pairs in O*NET and their share in course descriptions. For this exercise, I restrict the task space to only those pairs that appear in any O*NET task descriptions. I measure the similarity between each field-occupation pair by calculating the cosine distance in task space. To ensure that I focus on meaningful matches, I drop any field-occupation pair with either an empirical share below 0.01% in the ACS and no task overlap.[10] Finally, I compute the rank-rank correlation between the

[9] A robustness exercise that uses only early-career workers (ages 25-35) produces similar field-to-occupation mappings.

[10] Dropping these occupation-field pairs is generally conservative for this validation exercise, as it excludes

empirical field-to-occupation shares and the task-based rankings.

The resulting rank-rank correlation is 0.31, indicating a meaningful positive relationship between the two measures. Given the sparsity of O*NET task lists and the fact that some fields map to multiple occupations imperfectly, this relatively large and positive correlation validates the relevance of my task-based measure. It also provides new evidence of how the tasks students learn in college align with the jobs they enter after graduation.

5 Results

This section documents three facts about curricular exposure to AI. First, college coursework is substantially more exposed to AI capabilities than to earlier technology shocks. Second, exposure to AI has risen over time, driven by the spread of data analysis and related tasks across fields. Third, exposure varies meaningfully across institutions within the same broad field of study. I summarize descriptive findings on field-level exposure and assess distributional impacts of heterogeneous exposure across different groups.

5.1 Overall overlap between courses and technology

Figure 2 summarizes field-level exposure to three technologies using course descriptions from the 2023-2024 academic year. For each of 52 undergraduate fields of study, I pool verb-object pairs across institutions and calculate exposure to Robots, patent-based AI, and the LLM-based "OpenAI" measure. To facilitate comparison across technologies, I normalize each field-level score by the occupational exposure distribution for the same technology.

The first panel of Figure 2 shows that exposure to Robots is low across nearly all fields. The most exposed fields are vocational and applied fields such as Skilled Trades, with modest additional overlap in Engineering, Agriculture, and some technical STEM areas. Relative to the occupational distribution, most fields of study remain below the mean level of Robot exposure.

The next two panels show a very different pattern for AI. Both the patent-based AI measure and the "OpenAI" measure, which captures overlap between course tasks and LLM capabilities, imply much greater exposure than Robots. Between the two, there is greater variation in field-level exposure under the patent measure compared to the OpenAI measure. This is substantively useful: the patent-based AI measure captures overlap with the broader frontier of AI-related patented capabilities, while the OpenAI measure captures overlap with a narrower but more immediately relevant set of LLM capabilities. The patent-based

pairs where the occupation-to-field shares and task overlap are strongly correlated.

measure therefore shows more dispersion across fields; the OpenAI measure indicates that LLM exposure is widespread, even outside the most quantitative majors.

By normalizing according to each technology's occupation-level distribution, the figure asks where a field of study would fall relative to the distribution of occupations exposed to that technology. On this scale, Robot exposure in college coursework is generally low. AI exposure is much higher. In particular, the OpenAI measure indicates that a large share of fields cluster toward the upper part of the occupational exposure distribution, while the patent-based AI measure shows a longer right tail among highly quantitative fields.

Figure 3 presents the exposure distributions as a CDF, indicating where different fields of study fall in the exposure distributions. The first panel plots the CDF of field-level exposure to capabilities of Robots. The most-exposed field of study is Skilled Trades, indicating substantial overlap between tasks typical of Automotives, Construction, and other trades with Robots. Fields are typically less exposed to Robots than the average occupation.[11]

In both AI measures, Statistics/Data Science is at the top of the exposure distribution, thought the degree to which it is more exposed than the typical field of study varies by exposure measure.[12] Under the patent-based measure, STEM fields are highly exposed to AI, largely due to their emphasis on data analysis and related tasks. Under the OpenAI measure, which captures primarily a field of study's overlap with the capabilities of LLMs (based on the capabilities of GPT-4), Statistics and Computer Science are the most exposed fields — for their emphasis on coding and data analysis — followed by English, then Social Science fields that involve a mix of writing and data analysis skills. The OpenAI curve is also less steep than the patent-based AI curve, reflecting less cross-field dispersion and more common exposure across majors. Put differently, the patent-based AI measure is better at highlighting the extreme right tail of highly exposed fields, while the OpenAI measure better captures the broad diffusion of LLM-relevant capabilities across the curriculum.

The central takeaway from Figures 2 and 3 is therefore straightforward: college coursework is far more exposed to AI — and especially to LLM capabilities — than to earlier technology shocks. This pattern is consistent with the idea that AI differs from prior waves of automation in the extent to which it overlaps with high-skill, cognitively demanding tasks taught in college.

[11] It is notable that STEM fields have higher exposure than non-STEM fields, potentially indicating a limitation of the patent-based exposure scoring. Because patentable activities overlap disproportionately with STEM fields compared to other fields, some of the higher alignment likely reflects spurrious overlaps based on subject rather than task.

[12] Reassuringly, Statistician is approximately 2 occupation SDs more exposed to AI under the patent-based measure than the typical occupation, suggesting that the same overlap coming through in Statistics courses is connected to the tasks/skills involved in being a Statistician.

5.2 Time trend in technology exposure

Figure 4 shows how exposure to each technology has evolved over the past two decades. I estimate these trends from regressions of institution-field-year exposure on year indicators with institution-by-field fixed effects, so the figure isolates within-field changes over time rather than changes driven by shifting field composition.

The clearest pattern is the steady rise in patent-based AI exposure, while exposure according to the Robot and OpenAI measures is largely flat. On the patent-based AI exposure measure, over time, coursework increasingly incorporates tasks that overlap with the kinds of coding and data analysis capabilities now performed by AI tools. The timing of the increase is also informative. Exposure begins to rise noticeably in the early 2010s and continues to increase through the 2020s. This pattern suggests that curricular exposure to AI predates the public release of ChatGPT. Universities did not suddenly become exposed to AI in late 2022; rather, they had already been moving toward task content that overlaps with modern AI for more than a decade.

Substantively, the time trend appears to be driven by the spread of data analysis and related empirical tasks across fields. As shown in the validation section, tasks such as "analyze data" appear across a broad range of disciplines rather than only in Statistics or Computer Science. The time series evidence indicates that this diffusion has continued over time, raising AI exposure both within and outside traditionally quantitative fields.

By contrast, exposure to Robots remains low and essentially flat. The shift in coursework from more writing and analytical skills towards greater emphasis on data analysis appears as a shift within the patent-based AI exposure measure, but as these are commonly exposed skills under the OpenAI measure, the figure observes a modest or negligible increase in exposure over the last 20 years.

5.3 Heterogeneity by institution type

The aggregate field-level results mask substantial heterogeneity across institutions. To examine this heterogeneity, I calculate exposure separately for field-by-institution-type cells, grouping institutions into R1 universities, R2 universities, teaching-focused four-year institutions, and community colleges. I control for field fixed effects; thus, the interpretation of the estimates is differential exposure to a given technology class *for classes in the same field*.

Table **??** summarizes the distribution of exposure across these institution categories. The table documents substantial variation in exposure across institution types. On Robot exposure, courses at Community Colleges are significantly more Robot-exposed than classes in the same field offered at four-year institutions. The result is driven strongly by skilled

trade courses, which are substantially more vocational at Community Colleges and emphasize skills that overlap with the capabilities of Robots.

According to both AI exposure measures, R1 institutions offer courses that are more AI-exposed than at less research-oriented institutions. Differences between R1 institutions and Community Colleges are large and statistically significant under both measures; differences under the patent-based measure are large at other institution types and statistically different between R1s and teaching focused institutions, while differences are more modest under the OpenAI measure and significantly different only at Liberal Arts colleges and Community Colleges.

These differences appear to arise less from different treatment of the same course title than from differences in course menus and curricular depth. Teaching-focused institutions tend to offer fewer advanced electives and a larger share of intermediate, applied, and professionally oriented coursework. Research universities, by contrast, offer both very advanced theoretical courses and a wider range of specialized electives. Depending on the field, either structure can generate high exposure: teaching-focused institutions may concentrate more heavily on applied tasks already within reach of LLMs, while research universities may generate very high exposure in fields such as Statistics/Data Science because they offer a large volume of data-intensive electives.

This distinction matters for interpretation. The relevant heterogeneity is not primarily whether the "same" course is taught differently at different institutions, but whether students in the same major are exposed to different mixes of courses. In that sense, exposure is partly a feature of curricular composition. The results underscore a broader point of the paper: the same broad degree label can mask substantial differences in the task content of the coursework students actually complete.

The right tail of the institution-type distribution is dominated by quantitative fields, especially Statistics/Data Science. This is consistent with the broader descriptive results and suggests that the expansion of applied statistics and data science coursework has been an important driver of curricular exposure to AI.

5.4 Distributional implications

The preceding results imply that AI exposure is not evenly distributed across students. Because exposure is concentrated in fields linked to relatively high earnings, the rise of AI may compress some of the earnings differences that have historically favored quantitatively intensive majors. At the same time, the institutional heterogeneity documented above implies that students with the same nominal field of study may face meaningfully different exposure

depending on where they enroll.[13]

Figure 5 provides descriptive evidence on these distributional patterns using the OpenAI measure. The figure shows how exposure varies across the gender composition of degrees, across the earnings associated with majors, and across institution types. Three patterns are worth highlighting. First, highly exposed majors are disproportionately associated with higher-earning fields. Second, exposure is particularly concentrated only in elite research settings, though it is distributed across several parts of the four-year sector. Third, exposure is more acute among men than women due to gender sorting across majors, implying that changes in the returns to AI-exposed skills are likely to have uneven incidence across genders.

These descriptive patterns do not, on their own, establish welfare effects. Exposure remains agnostic about whether AI ultimately substitutes for or complements the skills taught in college. But they do indicate that any labor market consequences of AI are likely to be mediated through the field structure of higher education and through heterogeneity in the curricula offered across institutions.

6 Instructor and student responses to ChatGPT release

Having documented substantial exposure of college coursework to AI, this section examines how universities and students respond to a sharp, recent technology shock: the release of ChatGPT in November 2022. The rapid diffusion of LLM tools provides a natural setting to assess whether highly exposed coursework adjusts along either the intensive margin (course design) or extensive margin (course-taking behavior).

The release of ChatGPT has two immediate implications for higher education. First, it lowers the effective cost of completing coursework in tasks that overlap with AI capabilities, as students can use these tools to augment or substitute for their own effort. Second, it introduces uncertainty about the labor market returns to the skills developed in these courses. These forces generate ambiguous predictions for enrollment, but clearer implications for course design: if instructors seek to preserve the signal of student learning, they may shift evaluation toward forms of assessment that are less easily completed with AI assistance

[13]Throughout this the following section, I treat the LLM-based "OpenAI" measure as my preferred measure of AI exposure because it is most directly connected to the capabilities of the tools students and instructors are currently using. While the patent-based AI measure captures the broader frontier of patentable AI capabilities, the OpenAI measure is better aligned with the short-run question of how exposed coursework is to the specific kinds of LLM tools that students and instructors can already use. It also has a conceptual advantage over many existing AI exposure measures. In the OpenAI data, a task can be highly exposed because a single LLM capability substantially reduces the time required to complete it. Exposure therefore does not require overlap with many distinct AI capabilities. By contrast, patent-distribution approaches such as Webb's reward broader overlap across the distribution of capabilities, and ability-based approaches such as Felten et al. (2021) similarly place greater weight on tasks that map to multiple AI capabilities.

(e.g., in-class exams, participation) and away from assignments that can be completed out of class (e.g., essays, homework).

I evaluate these predictions using an event study framework around the release of ChatGPT, focusing first on instructor responses and then on student course-taking behavior.

6.1 Adaptation of course design

To assess instructor responses, I use a panel of syllabi from 20 institutions spanning the pre- and post-ChatGPT periods.

To document that instructors were cognizant of ChatGPT's release, and concerned about its impact on learning production, Figure 6 plots estimates from an interrupted time series regression capturing the increase in share of syllabi with an AI policy. Observations in this regression are at the course-semester-year level, and the regression controls for course fixed effects, thereby tracking within-course changes over time. The figure shows a sharp increase in the share of syllabi that explicitly reference AI following the release of ChatGPT. Prior to Fall 2022, fewer than 1% of courses include any mention of AI tools. By Fall 2025, more than half of syllabi include an AI policy. AI policy adoption was greatest in Social Science and Humanities courses, and most policies imposed restrictions or regulations on student use. The pattern indicates that instructors are aware of and concerned about the implications of AI, even as changes to course structure remain limited.

Next, I study how instructors adapted their courses to a post-ChatGPT paradigm. From each syllabus, I use an LLM to extract the weight assigned to major grade components (exams, homework, essays, participation, presentations, and projects).[14] I estimate event study regressions of the form:

$$y^q_{i,s,t} = \alpha^q_{i,s} + \sum_{t \in T} \beta^q_t + \varepsilon^q_{i,s,t},$$

where $y^q_{i,s,t}$ is the weight on grade component q in course i at institution s in semester t, and $\alpha^q_{i,s}$ are course fixed effects. The coefficients β^q_t trace out changes in grading weights relative to Fall 2022, the last pre-ChatGPT term.

Figure 7 presents the results. The most striking feature of the figure is how little course structure changes following the release of ChatGPT. There is a gradual decline in the weight placed on essays and modest increases in participation and presentation components, consistent with instructors shifting toward more in-class forms of assessment. However, these

[14] Human validation of the LLM extraction suggests accuracy above 85%. The most common error in extraction is to miss the assessments altogether; hallucinations or errors in extraction represent less than 5% of cases.

changes are small in magnitude and evolve slowly over multiple semesters. To contextualize these effects, by Fall 2025 the decline in essay weight corresponds to a 11% change relative to the Fall 2022 baseline, while increases in participation and presentation account for a change. These adjustments are economically modest relative to baseline variation in grading schemes across fields and institutions.

The limited response to ChatGPT stands in sharp contrast to the large and immediate changes in course structure observed during the COVID-19 pandemic. Beginning in Spring 2020, courses shifted rapidly away from exams and toward homework-based and take-home assessments, with effects that persisted well beyond the initial shock.[15] Relative to this earlier episode, the response to ChatGPT is both smaller in magnitude and more gradual, suggesting that instructors face different constraints when adapting to technologies that affect the production and evaluation of student work.

Taken together, these results suggest that instructors respond quickly along low-cost margins, such as adding policies, but adjust more slowly along higher-cost margins, such as redesigning course evaluation and structure.

6.2 Course taking

I next analyze whether students adjust their course-taking behavior in response to the release of ChatGPT. The release of ChatGPT has, ex ante, ambiguous effects on students' incentives to enroll in AI-exposed courses. High AI exposure reduces the cost of taking a course, but may affect the expected return to developing skills in these courses. The theoretical effect depends on the direction of students' expectations for the earnings impacts of AI exposure.

To evaluate these competing forces, I estimate event study models of enrollment by field and course-level AI exposure, comparing changes in enrollment patterns before and after Fall 2022. Using the course catalog data, I aggregate (log) enrollment and course/section counts to the institution-field-semester-year level and estimate separate interrupted time series regressions of log enrollment or course supply on year dummies interacted with field-level AI exposure,[16] controlling for institution-field, institution-semester, field-semester, and institution-year fixed effects.

Preliminary results indicate limited changes in enrollment patterns following the release of ChatGPT. Across fields and exposure levels, enrollment trends remain largely continuous through the post-ChatGPT period, with no evidence of sharp reallocation toward or away

[15] Given that the majority of schools in my syllabus sample are in Texas, which offered in-person instruction in Fall 2021, the extent of these changes are likely to underestimate the national responses to pandemic learning constraints.

[16] Using the OpenAI exposure measure. Results are substantively similar using the patent-based AI exposure measure.

from highly exposed coursework. These findings mirror the limited adjustments observed on the supply side. Despite the rapid adoption and widespread availability of AI tools, both instructors and students exhibit relatively muted behavioral responses in the short run.

6.3 Discussion

The contrast between large measured exposure to AI (Section 5) and limited observed adjustment highlights an important feature of higher education: curricular and behavioral responses to technological change appear slow-moving, even when the underlying technology diffuses rapidly. The comparison to the pandemic period is particularly informative. While universities adapted quickly to constraints on in-person instruction, the response to AI — which alters the production and evaluation of human capital rather than the feasibility of instruction itself — has been more gradual and limited.

These patterns suggest that institutional frictions, uncertainty about the long-run role of AI, and the difficulty of redesigning courses may constrain the short-run responsiveness of higher education to technological change.

7 Conclusion

This paper adapts the approach of Webb (2019) to develop new measures of the exposure of college curricula to emerging technologies, including Artificial Intelligence. Whereas Webb (2019) map patents to O*NET tasks to identify which jobs are technologically exposed, I map these technologies to the tasks embedded in college courses and majors. This shift moves the focus from labor demand adjustment to upstream investments in human capital and the extent to which educational content overlaps with evolving technological capabilities.

The descriptive results confirm predictions in the literature that AI is distinct from earlier technological waves in the degree to which it overlaps with high-skill tasks. Exposure of college courses to Robots is generally low, but the median field of study has AI exposure comparable to the 75th percentile of the occupation-level AI exposure distribution. STEM and Business/Economics courses are the most exposed to AI, as are courses at teaching-focused institutions relative to research universities.

Taken together, these findings suggest that AI capabilities already overlap substantially with the content of postsecondary education, including in fields traditionally associated with high-skill human capital formation. At the same time, complementary analyses in the paper document relatively modest adjustments in course offerings and enrollment patterns following the recent AI shock, pointing to a potential disconnect between the rapid evolution of technological capabilities and the slower-moving structure of higher education.

More broadly, an open question is whether (and in what ways) task-technology overlap reflects a risk of substitution versus an opportunity for productivity gains. I remain intentionally agnostic about whether exposure implies a threat of automation or a form of complementarity that enhances human capabilities. Resolving this distinction is central for interpreting the policy implications of these findings and represents an important direction for future research.

Works cited

Acemoglu, Daron et al. (2022). "Artificial intelligence and jobs: Evidence from online vacancies". In: *Journal of Labor Economics* 40.S1, S293–S340.

Autor, David H, Lawrence F Katz, and Melissa S Kearney (2006). "The polarization of the US labor market". In: *American economic review* 96.2, pp. 189–194.

Autor, David H, Frank Levy, and Richard J Murnane (2003). "The skill content of recent technological change: An empirical exploration". In: *The Quarterly journal of economics* 118.4, pp. 1279–1333.

Biasi, Barbara and Song Ma (2022). *The education-innovation gap*. Tech. rep. National Bureau of Economic Research.

Brynjolfsson, Erik and Tom Mitchell (2017). "What can machine learning do? Workforce implications". In: *Science* 358.6370, pp. 1530–1534.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock (2018). "What can machines learn and what does it mean for occupations and the economy?" In: *AEA papers and proceedings*. Vol. 108. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 43–47.

Desell, Travis, James Higgins, and Sophine Clachar (2019). *Analyzing flight data using predictive models*. US Patent 10,248,742.

Eloundou, Tyna et al. (2023). "Gpts are gpts: An early look at the labor market impact potential of large language models". In: *arXiv preprint arXiv:2303.10130* 10.

Felten, Edward, Manav Raj, and Robert Seamans (2021). "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses". In: *Strategic Management Journal* 42.12, pp. 2195–2217.

Honnibal, Matthew et al. (2020). "spaCy: Industrial-strength natural language processing in python". In.

Javadian Sabet, Alireza et al. (2024). "Course-Skill Atlas: A national longitudinal dataset of skills taught in US higher education curricula". In: *Scientific Data* 11.1, p. 1086.

Jurafsky, Dan and James H Martin (2014). *Speech and language processing. Vol. 3*.

Light, Jacob (2026). "Student demand and the supply of college courses". In: *Available at SSRN 4856488*.

McElheran, Kristina et al. (2024). "AI adoption in America: Who, what, and where". In: *Journal of Economics & Management Strategy* 33.2, pp. 375–415.

Timmerman, Jean Xiao (2025). "Educational Exposure to Generative Artificial Intelligence". In: *FEDS Notes* 2025-02, pp. 26–2.

Webb, Michael (2019). "The impact of artificial intelligence on the labor market". In: *Available at SSRN 3482150*.

Figures and Tables

Figure 1. Example course



ECON 101
Lecture: 01
Units: 5
Class#: 7559

Winter 2026
Open

Economic Policy Seminar

Department of Economics

Lecture, Discussion
1/5/26 - 3/13/26

Mon, Wed
1:30 PM - 3:20 PM
Econ 140

Instructor: Light, J.

Enrollment Status

Open Seats: 7
Enrolled: 17 Waitlist: 0
Capacity: 24 Waitlist Max: No waitlist

Course Description

Capstone and writing in the major course open to Econ majors only. Economic policy analysis, writing and oral presentations will be large components of this course. Students may also complete group projects that include empirical economic analysis focused on a specific topic. The goal of this course is to enable students to utilize the skills they have acquired throughout their time in the major. Section topics vary by instructor. Enrollment limited to application at the start of each school year with student placement notifications before the term starts. Permission numbers will be provided to students. Limited to students applying to graduate in 2025-26. Enrollment by application: https://economics.stanford.edu/forms.

Grading basis
Letter (ABCD/NP)

Class level
Undergraduate

Instructional mode
In Person

Final exam
Wed, March 18th
3:30 PM - 6:30 PM
Econ 140

Gen Ed Requirement(s)
WAYS - Social Inquiry (SI)

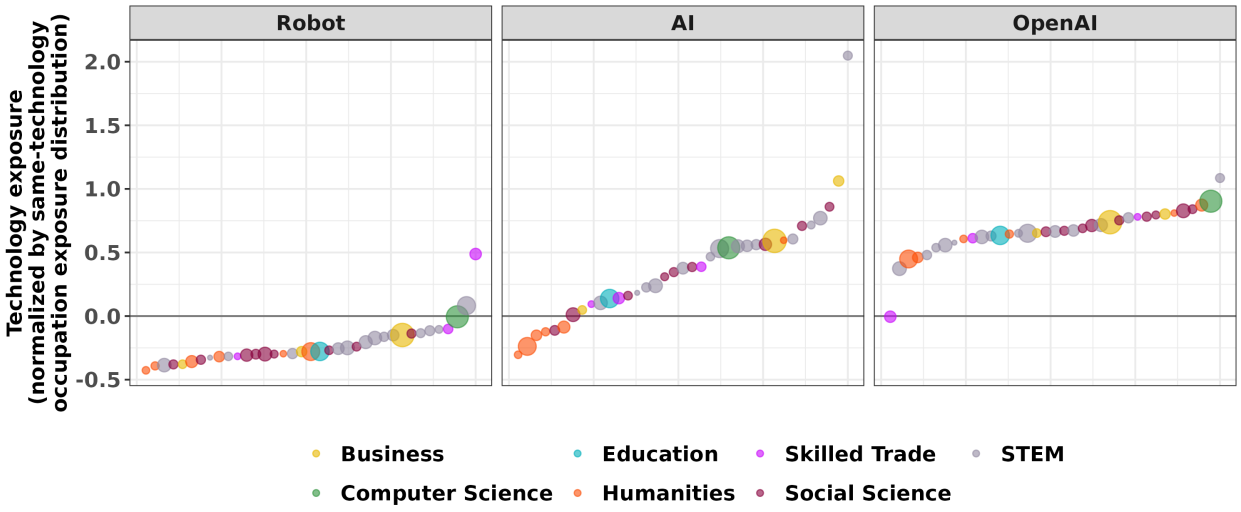
Source: Stanford University.

Figure 2. Technology exposure distribution across fields



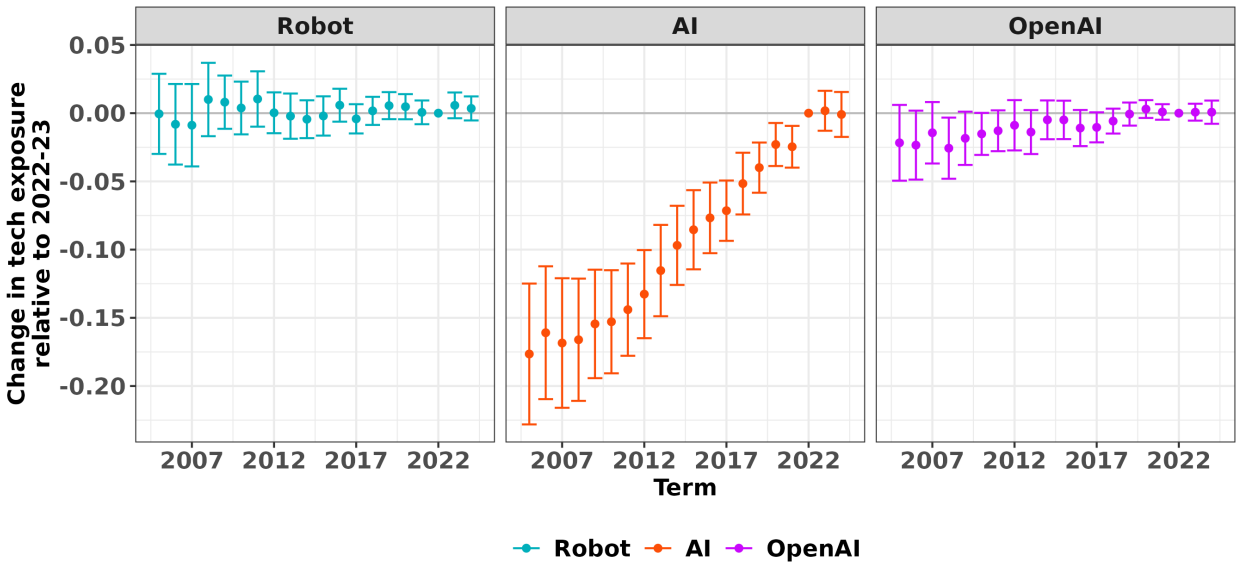
Notes: The figure plots density plots of field-technology exposure scores, split by technology class. "Robot" and "AI" measure patent-based overlap using the Webb (2019) exposure methodology; the "OpenAI" score measures overlap based on LLM capabilities measured by Eloundou et al. (2023). Courses are partitioned into 52 fields (e.g., Math, Business, History). Scores are normalized based on the occupation distribution of the same technology.

Figure 3. Technology exposure trend across fields - CDF



Notes: The figure plots density plots of field-technology exposure scores, split by technology class. "Robot" and "AI" measure patent-based overlap using the Webb (2019) exposure methodology; the "OpenAI" score measures overlap based on LLM capabilities measured by Eloundou et al. (2023). Courses are partitioned into 52 fields (e.g., Math, Business, History). Each field of study is represented as a dot, with size corresponding to the enrollment share in the field.

Figure 4. Time trend in technology-field of study exposure



Notes: The figure plots density plots of field-technology exposure scores, split by technology class. "Robot" and "AI" measure patent-based overlap using the Webb (2019) exposure methodology; the "OpenAI" score measures overlap based on LLM capabilities measured by Eloundou et al. (2023). Courses are partitioned into 52 fields (e.g., Math, Business, History). Observations are at the institution-field-year level, estimates from a time series regression, with 2022-23 the reference year, controlling for institution-by-field fixed effects and weighting by the number of verb-object pairs in each cell. Scores are normalized based on the occupation distribution of the same technology.

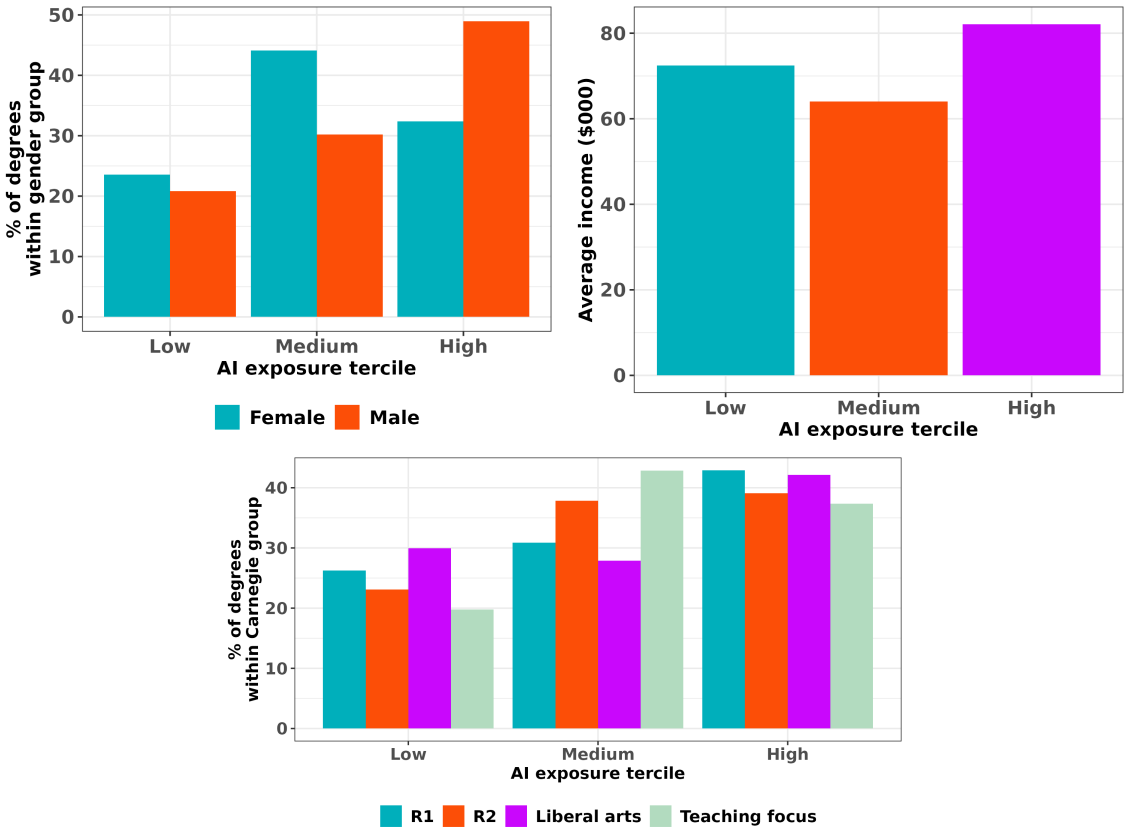
Table 1. Technology exposure by institution category

Dependent Variables:	Robot	AI	OpenAI
Model:	(1)	(2)	(3)
Variables			
R2	0.0024	-0.0347	-0.0095
	(0.0192)	(0.0473)	(0.0197)
Liberal arts	0.0120	-0.0565	-0.0644***
	(0.0294)	(0.0593)	(0.0240)
Teaching focus	0.0081	-0.0895**	0.0032
	(0.0170)	(0.0431)	(0.0169)
Community college	0.1534***	-0.2709***	-0.1573***
	(0.0428)	(0.0573)	(0.0278)
Fixed-effects			
Field	Yes	Yes	Yes
Fit statistics			
Observations	11,288	11,288	12,974
R^2	0.08439	0.08247	0.12558

Clustered (institution) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

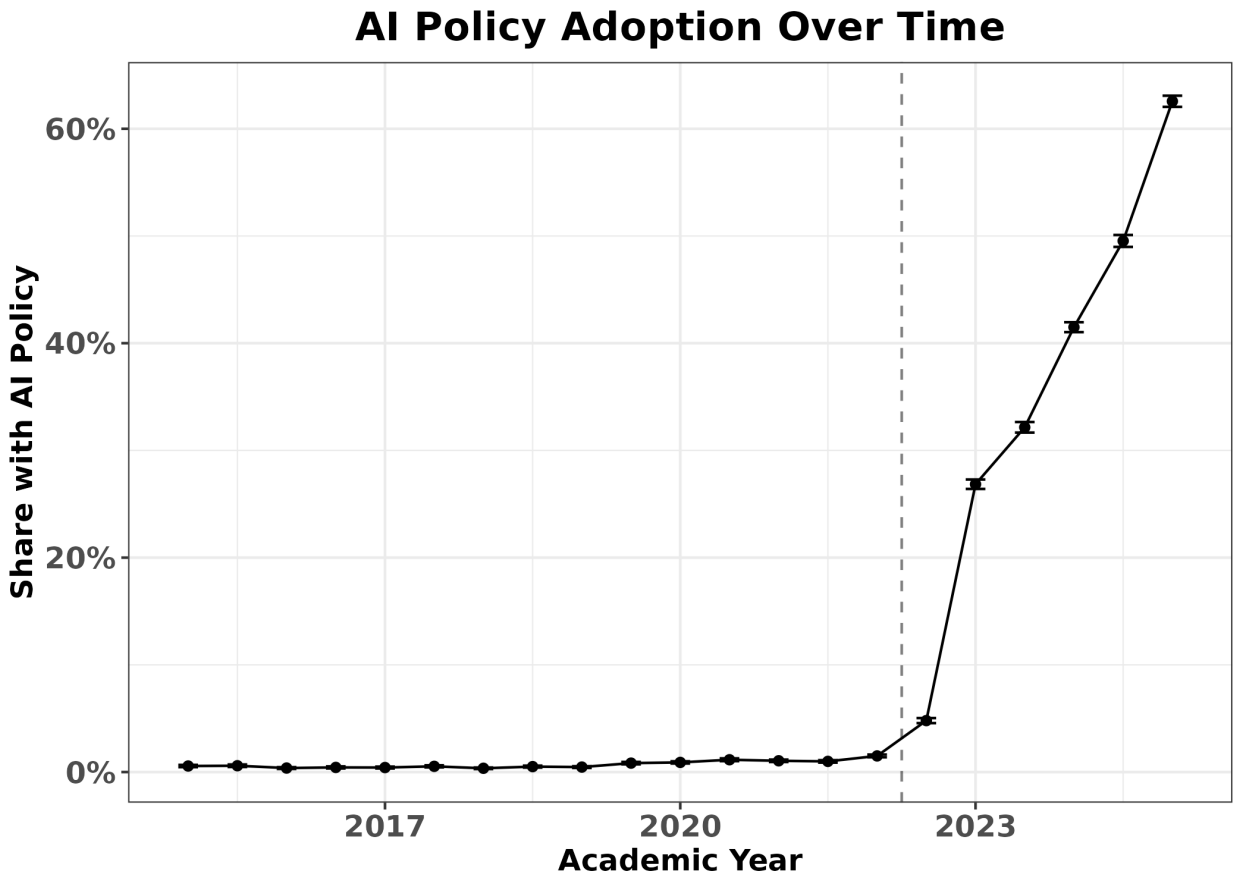
Notes: The table summarizes coefficients from separate regressions of field of study exposure on indicators for institution type. In each regression, observations are at the institution-field level and control for field of study fixed effects and weighted by the number of verb-object pairs in each cell. The reference category in each regression is R1 (research) universities. Standard errors are clustered at the institution level.

Figure 5. Distributional implications of technology exposure



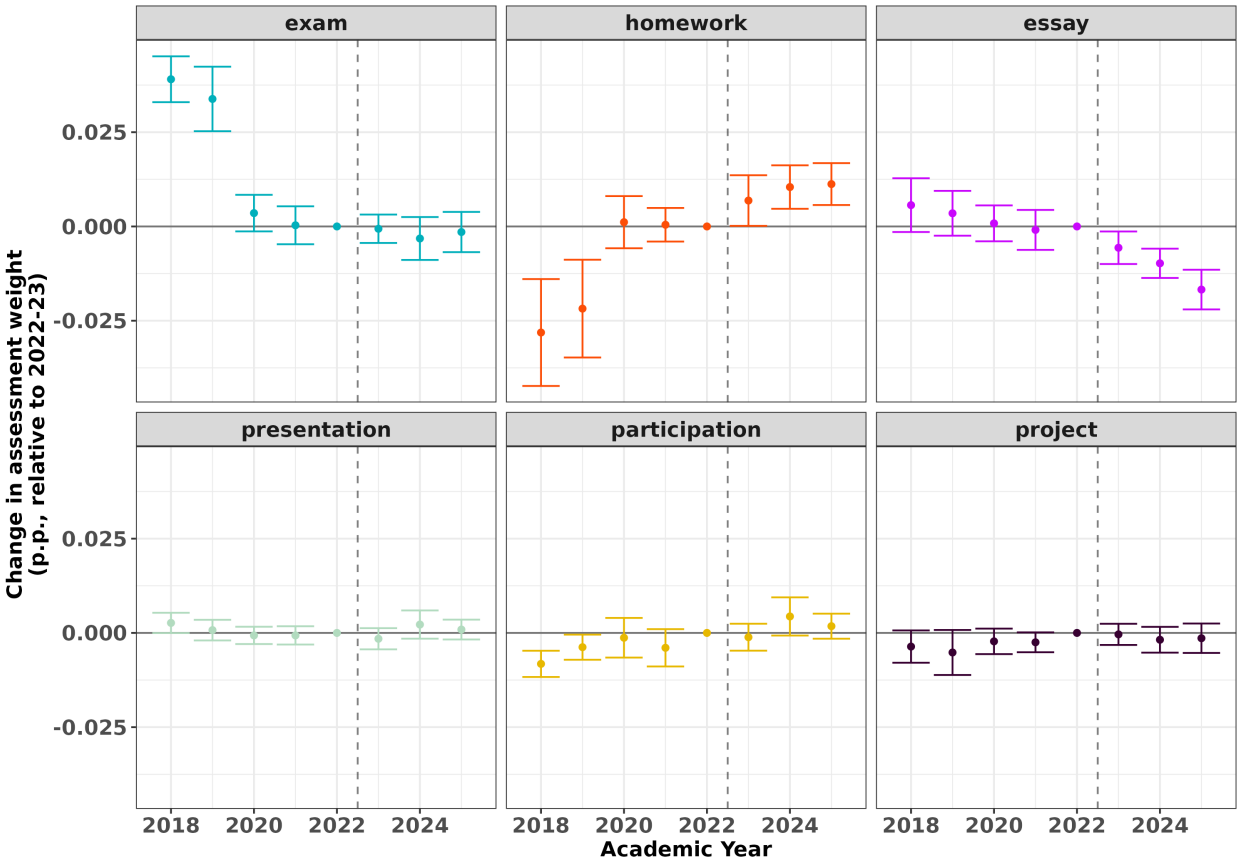
Notes: The three figures summarize differential technology exposure across three different groups, according to terciles of the "OpenAI" LLM-based AI exposure distributions. The first panel plots differential exposure for men and women based on the within-gender share of major completions at four-year institutions. The second panel plots average income ($000) by tercile of AI exposure. The third panel plots the within-Carnegie group share of major completions.

Figure 6. Trend in AI policies



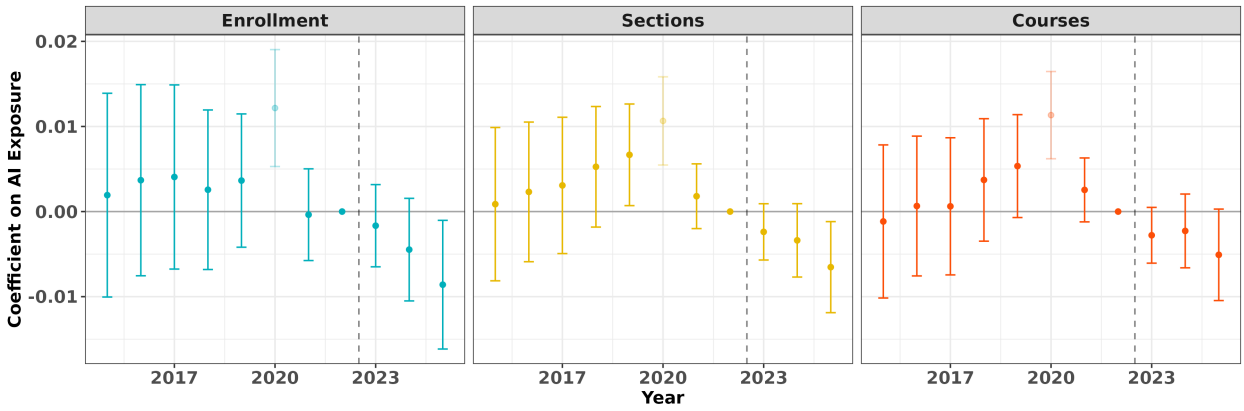
Notes: Estimates from interrupted time series with binary dependent variable (1 if the syllabus has an explicit AI policy), based on course-level observations from a sample of syllabi from 20 US colleges and universities. Fall 2022 is the omitted term. Time variable in academic years (thus, the 2022 point refers to the Fall 2022 semester, the 2022.5 point refers to Spring 2023, etc.). The dashed line indicates the public release of ChatGPT. AI policies pre-Fall 2022 are typically policies regarding the use of online translation tools in foreign language courses or related to AI-assisted exam administration during the pandemic. Standard errors clustered at the course level.

Figure 7. Changes to assessment weights around release of ChatGPT



Notes: Grading weights from syllabi from a sample of 20 US colleges and universities. Estimates come from interrupted time series regressions of each assessment weight category around the release of ChatGPT in 2022-23. Observations are at the institution-course-semester level. Regressions control for course-semester fixed effects. Standard errors clustered at the course level.

Figure 8. Enrollment trend following ChatGPT release



Notes: Estimates come from interrupted time series regressions of log course enrollment, and log quantity of course and section offerings, around the release of ChatGPT in 2022-23. Observations are at the institution-term-semester-field of study level. Regressions control for institution-field, institution-semester, field-semester, and institution-term fixed effects. Standard errors clustered at the institution-field level.

Table 2. Common verb-object pairs for select fields

Field (1)	Verb (2)	Object (3)	% of tasks (4)	AI (5)	Robot (6)	Software (7)
Arts	develop	technique, concept, approach, appreciation	0.023	0.000	0.000	0.000
Arts	create	work, project, artwork, image	0.023	0.004	0.002	0.006
Arts	teach	music, lesson, skill, technique	0.018	0.000	0.000	0.001
Chemistry	perform	experiment, calculation, analysis, research	0.067	0.015	0.004	0.012
Chemistry	develop	technique, concept, perspective, principle	0.038	0.000	0.000	0.000
Chemistry	qualitative	analysis	0.035	0.000	0.000	0.000
Computer Science	develop	application, solution, system, program	0.019	0.003	0.000	0.020
Computer Science	process	spreadsheet, database, communication, graphic	0.019	0.000	0.000	0.003
Computer Science	create	application, page, visualization, plan	0.010	0.007	0.008	0.014
Economics	analyze	datum, problem, behavior, issue	0.045	0.124	0.028	0.050
Economics	develop	model, tool, theory, technique	0.033	0.019	0.004	0.003
Economics	apply	tool, method, analysis, model	0.032	0.038	0.009	0.006
Education	teach	lesson, skill, art, child	0.025	0.000	0.000	0.001
Education	learn	strategy, environment, disability, style	0.022	0.013	0.000	0.000
Education	develop	proposal, relationship, approach, philosophy	0.012	0.000	0.000	0.000
English	write	composition, requirement, analysis, fiction	0.063	0.000	0.000	0.000
English	read	text, literature, page, skill	0.034	0.000	0.000	0.001
English	develop	process, expertise, paper, question	0.023	0.000	0.000	0.001
Ethnic/Cultural Studies	examine	relationship, aspect, method, content	0.023	0.000	0.000	0.000
Ethnic/Cultural Studies	analyze	impact, representation, issue, study	0.021	0.003	0.000	0.002
Ethnic/Cultural Studies	develop	tool, appreciation, proposal, sense	0.016	0.004	0.000	0.001
Math	solve	equation, system, variety, model	0.194	0.000	0.002	0.005
Math	analyze	datum, graph, problem, relationship	0.028	0.091	0.021	0.038
Math	graph	system, analysis	0.025	0.007	0.000	0.000
Stats/Data Science	analyze	datum, output, set, result	0.078	0.214	0.052	0.088
Stats/Data Science	present	datum, result, concept, model	0.040	0.015	0.000	0.025
Stats/Data Science	test	analysis, power, model, application	0.017	0.001	0.000	0.002

Notes: The table lists the three most common verbs in each field of study and, for each verb, the (up to) four most common associated objects. Column 4 reports the share of all verb-object pairs represented by these selected pairs. Columns 5-7 multiply these shares by the technology weights for each technology class.