The Content of Opportunity:

Community College Courses and Student Outcomes*

Barbara Biasi† Song Ma‡ Zhengren Zhu§

March 24, 2026

Preliminary and incomplete. Please do not cite without the authors' permission.

Abstract

Community colleges enroll nearly half of all U.S. undergraduates, yet we know strikingly little about what is taught in these institutions and how this shapes students' economic trajectories. Using 605,913 syllabi from 40 Texas community colleges, we measure course content along two dimensions that capture the dual mission of community colleges: similarity to four-year curricula (academic relevance) and coverage of locally in-demand labor market skills (labor-market relevance). Linking syllabi to administrative records and exploiting last-minute syllabus updates unobservable to students at enrollment, we find sharply divergent effects of course content across content dimensions and student pathways. For students in academic programs, higher academic relevance raises transfer, degree completion, and earnings; higher labor-market relevance reduces degree attainment, consistent with misalignment between vocational content and academic goals. The earnings effects of specific skill families also differ across student pathways. Skills that raise earnings for academic students tend to be analytical (e.g., law and analysis), while those for vocational students tend to be service- and business-oriented (e.g., service and administration). Conversely, many analytical skills (e.g., law) reduce vocational students' earnings, while many service- and business-oriented skills (e.g., service) decrease academic students' earnings. We conclude that course content — and its alignment with student pathways — is a largely overlooked dimension of community college policy, with misaligned content potentially generating negative returns.

JEL Classification: I23, I24, I26, J24, J31

Keywords: Community Colleges, Course Content, Skills, Instructors, Economic Mobility, Inequality

*The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. All errors are our own.

†Yale School of Management and NBER, barbara.biasi@yale.edu;

‡Yale School of Management and NBER, song.ma@yale.edu.

§Vassar College, zhengren.zhu@gmail.com.

1 Introduction

Community colleges are a central pillar of U.S. higher education and, in principle, an important vehicle for economic mobility. Charging tuition below one-third of public four-year universities and admitting any applicant with a high school diploma, they enroll approximately 10.5 million students each year — roughly 39 percent of all undergraduates — and are the modal entry point into post-secondary education for first-generation and low-income students (Bailey et al., 2015; American Association of Community Colleges, 2023). Community colleges serve a dual mission: they offer academic programs designed to transfer students into four-year institutions, and vocational programs designed to prepare students for direct labor market entry (Cohen et al., 2014). Whether they succeed at these goals — and what determines whether they do — are questions with direct implications for educational inequality and economic opportunity.

Yet community colleges are often perceived as lower quality than four-year institutions, and existing evidence supports some of this concern. Students who enter four-year institutions through community college transfer face lower bachelor's degree completion rates than those who enroll directly (Rouse, 1995; Kane and Rouse, 1995; Mountjoy, 2022). The returns to community college attendance are positive on average but smaller and more heterogeneous than returns to four-year schooling, and vary substantially with the type of program and credential (Kane and Rouse, 1995; Jepsen et al., 2014; Stevens et al., 2019; Grosz, 2020). What remains largely unknown is *why*. In particular, we know strikingly little about the most basic input to the educational process in community colleges: what is taught, by whom, and whether these features of instruction have a causal effect on the outcomes community colleges are meant to deliver.

In this paper, we address these questions using a new dataset of 605,913 course syllabi from 40 Texas community colleges, linked to administrative records on students, instructors, and labor market outcomes. We measure course content along two dimensions that capture the dual mission of community colleges: *academic relevance* (AR), defined as the textual similarity of a community college course syllabus to its four-year counterpart, constructed using cosine similarity of TF-BIDF-weighted term vectors; and *labor-market relevance* (LMR), defined as a locally-weighted density of job-market skills in the syllabus, where weights are given by the demand for each skill cluster in the local labor market (Deming and Kahn, 2018). We study how these measures vary across institutions, fields, and instructors, and estimate their causal effects on student outcomes. We find substantially different effects across dimensions of content and student pathways. Academic alignment raises degree attainment and earnings for academic students, while alignment with labor market skills

reduces these outcomes. Labor-market skills have heterogeneous effects that depend on student pathways and the specific skills emphasized. Many skills that increase earnings for academic students reduce earnings for vocational students, and vice versa, highlighting the costs of misaligned course content.

Applying our measures to Texas syllabi reveals that community college courses are, on average, at least as similar to four-year courses as four-year courses are to each other — a finding that runs counter to the perception of a large quality gap. At the same time, the distribution is wide, and the sources of variation differ sharply between our two measures. For AR, most variation occurs across fields and persists within courses over time, while institutions and instructors play a smaller role. We do, however, find that instructors with doctoral degrees and prior experience at four-year universities are associated with higher academic relevance, consistent with faculty training shaping what is taught. For LMR among academic courses, the pattern is different: the college where a course is taught explains the majority of the variation, pointing to institution-level differences in curricular orientation rather than instructor discretion.

Estimating the causal effects of course content requires addressing a fundamental selection problem. Students are not randomly assigned to courses: those who are more academically motivated or better prepared are more likely to enroll in courses with higher AR, and would tend to have better educational outcomes regardless of course content. At the institution level, colleges that serve less advantaged populations may systematically differ in their curricular orientation. A naive comparison of outcomes across content levels would conflate the effect of course content with the effect of student and institutional sorting, producing biased estimates of either sign. We address this problem using two complementary strategies. Our primary strategy exploits an information friction in the course-registration process: students enroll based on the prior term's syllabus and cannot observe last-minute content updates made before the current semester begins. Because these within-course changes are unobservable at enrollment, they provide plausibly exogenous variation in both academic and labor-market relevance, conditional on lagged content. As a second strategy, we instrument for each student's average academic relevance using the content of courses offered at the geographically closest community college, exploiting the strong tendency of community college students to enroll near where they live.

AR has large positive effects on educational and earnings outcomes of students in academic programs. A one-standard-deviation higher AR raises the probability of earning an associate degree by 1.1 percentage points (relative to a baseline of 18.6%), of transferring to a four-year institution by

0.9 percentage points, and of completing a bachelor's degree by 1.5 percentage points; it also raises post-graduation earnings by 1.5 log points. These gains are broadly distributed across student groups but are somewhat larger for better-prepared and more financially advantaged students, likely because disadvantaged students face non-academic barriers that better coursework alone cannot overcome.

LMR has opposing effects for academic-program students, and its effects on vocational students depend critically on field. For academic students, a one-standard-deviation increase in LMR reduces associate degree completion by 3.6 percentage points and bachelor's degree attainment by 4.2 percentage points, consistent with vocational skill content crowding out the academic preparation that transfer-track students need. For vocational students, the average LMR effect on earnings is near zero, but this masks large heterogeneity across fields: LMR raises earnings by approximately 11 log points for students in Humanities and Other fields but reduces them by approximately 22 log points for students in STEM, where broader foundational preparation appears to matter more than narrowly targeted skill content.

The skill-level analysis further sharpens this picture and emphasizes the importance of skill alignment with student types. Decomposing LMR into its 29 constituent Lightcast skill clusters, we find that skills that raise earnings for academic students tend to be analytical (e.g., law, analysis, finance, and industry), while those for vocational students tend to be service- and business-oriented (e.g., service, administration, and HR). Conversely, many analytical skills (e.g., law, IT, and science) reduce vocational students' earnings, while many service- and business-oriented skills (e.g., service, security, and marketing) decrease academic students' earnings.

Taken together, the findings paint a nuanced picture. Content varies widely across colleges, and a non-trivial share of this variation traces back to who teaches the course: faculty with stronger academic credentials and closer ties to four-year institutions consistently produce more academically aligned syllabi. This variation matters, but its effects are not uniform. Academic alignment is unambiguously beneficial for students on the transfer track, but alignment with labor market skills is not uniformly good or bad; what matters is whether a course's orientation matches the trajectory of the student taking it. Within the broad categories of academic and labor-market relevance, the specific skill clusters emphasized can matter as much as aggregate indices suggest. In several cases, content that raises earnings for academic students reduces earnings for vocational students, and vice versa. In short, curricula design and its alignment with student objectives are consequential but underappreciated determinants of whether community colleges can deliver on their promises.

Related literature. This paper contributes to three strands of the literature. First, it contributes to the extensive literature on the returns to community college education and the role of community colleges as engines of economic mobility. A rich body of evidence establishes that community college attendance raises earnings and educational attainment, but that returns are smaller and more heterogeneous than at four-year institutions and vary substantially by program type and credential (Kane and Rouse, 1995; Rouse, 1995; Jepsen et al., 2014; Stevens et al., 2019; Grosz, 2020; Jacobson et al., 2005). Community colleges serve a disproportionately large share of low-income and first-generation students, making them a central institution in the study of intergenerational mobility (Deming and Dynarski, 2009; Chetty et al., 2020; Mountjoy, 2022). We depart from this literature by asking not whether community college attendance matters, but what features of course content determine how much it matters. To our knowledge, this is the first paper to measure course content at scale across a large community college system and to estimate causal effects of content on a broad range of student outcomes.

Second, this paper contributes to a growing literature on the content of instruction and its relationship to labor-market outcomes. A large body of work documents how skill demands have evolved — toward cognitive, technical, and interpersonal skills and away from routine tasks — using job postings, wage data, and occupational surveys (Autor et al., 2003; Acemoglu and Autor, 2011; Deming and Kahn, 2018; Deming, 2017; Hershbein and Kahn, 2018). A more recent literature studies educational content directly. Biasi and Ma (2026) measure the distance between university syllabi and the research frontier and find effects on student careers; Chau et al. (2023) link course content to workplace activities and earnings; and Biasi and Ma (2022) document a gap between what universities teach and the frontier of knowledge. On the economics of educational content more broadly, Altonji et al. (2012) document large heterogeneity in the returns to college major choices, highlighting that the specific content of a credential — not just its level — matters for labor-market outcomes. We complement this literature by constructing content measures at the course level for community colleges, a population underrepresented in prior work. We also find that instructor credentials and experience at four-year universities systematically predict academic alignment, consistent with evidence that instructor characteristics shape student outcomes through the quality of what is taught (Carrell and West, 2010; Bettinger and Long, 2010; Figlio et al., 2015).

Third, our results connect to the literature on vocational education, academic tracking, and student-curriculum fit. International evidence suggests that the returns to vocational relative to

general education depend on labor-market structure and the phase of the economic cycle, with vocational programs offering short-run earnings advantages that can erode over the career (Hanushek et al., 2017; Brunello and Checchi, 2007). Within the United States, a growing body of work documents the large and heterogeneous returns to career and technical education programs at community colleges (Jepsen et al., 2014; Stevens et al., 2019; Grosz, 2020). Our contribution to this strand is to show that the consequences of vocational content depend critically on who takes what courses: emphasis on certain specific labor-market skill that helps students in some program types and fields can be sharply harmful for others. This heterogeneity suggests that alignment between curriculum and student goals — rather than any single orientation being uniformly better — is a first-order determinant of educational value.

2 Institutional Setting

2.1 Community Colleges in the United States

Community colleges are largely public, predominantly two-year institutions that provide open-access postsecondary education with a dual mission: (i) academic programs that offer terminal associate's degrees and facilitate transfer to four-year institutions and (ii) vocational and technical training aligned with local labor markets (Cohen et al., 2014; Bailey et al., 2015). With approximately 6-7 million students annually (35-40% of all undergraduates in the U.S.), they are a central component of the country's higher education system (American Association of Community Colleges, 2023; National Center for Education Statistics, 2023).

Community colleges are cheap and non-selective. They admit any applicant with a high school diploma or GED, and many enroll students without traditional academic credentials. Tuition levels are low, typically below one-third of those at public universities (College Board, 2023). This pricing differential, combined with open admissions, lowers entry barriers relative to selective four-year institutions, and generates a student population that is older, more economically disadvantaged, and more racially and ethnically diverse than that of four-year colleges (Bailey et al., 2015; National Center for Education Statistics, 2023). This also makes community colleges the modal entry point into higher education for first-generation and low-income students (Deming and Dynarski, 2009).

Fields and programs. The curricular scope of community colleges spans general education fields (mathematics, natural sciences, social sciences, humanities) as well as highly specialized occupational coursework. Relative to four-year institutions, community colleges devote a larger share of instructional activity to career and technical education and to developmental education (Bailey

et al., 2015).

Community colleges offer three broad categories of programs:

1. *Academic transfer programs*, culminating in Associate of Arts (AA) or Associate of Science (AS) degrees. These are designed to be comparable to the curricula of the first two years of four-year institutions, often under statewide transfer agreements (Doyle, 2009).
2. *Applied and technical programs*, leading to Associate of Applied Science (AAS) degrees in fields such as nursing, information technology, advanced manufacturing, and business technology.
3. *Short-term certificates and workforce credentials*, frequently developed in collaboration with local employers and oriented toward rapid labor market entry.

Students enroll in community colleges for heterogeneous reasons. Some intend to transfer to a four-year institution after completing general education requirements; others seek occupation-specific training with immediate labor market payoff; still others enroll to upgrade skills, change careers, or re-enter education after a period out of school (Bailey et al., 2015).

2.2 The Texas Context

Texas operates one of the largest community college systems in the country, enrolling over 600,000 credit students annually across 50 public community college districts (Texas Higher Education Coordinating Board, 2023b). Governance is decentralized: community colleges are organized into taxing districts with elected boards, and funding combines state allocations, local property tax revenue (for districts with taxing authority), tuition, and federal funds (Texas Higher Education Coordinating Board, 2023b).

Open admissions is the norm in Texas, and tuition remains low relative to four-year public universities in the state, especially for in-district students (College Board, 2023). Because community colleges often serve geographically defined districts, in-district tuition rates are typically substantially below out-of-district or out-of-state rates, reinforcing the importance of local access.

State policy emphasizes articulation agreements to facilitate transfer to public universities, and Texas has implemented both common course numbering and a statewide core curriculum that attempts to standardize lower-division academic coursework across institutions. As a result, students across community colleges enroll in courses that are formally comparable in title, credit value, and learning objectives, even though instructional content and materials are determined locally. This combination of statewide standardization and local discretion plays a central role in our setting and motivates the institutional discussion that follows. At the same time, Texas community colleges

play a central role in workforce development, particularly in health professions, energy, logistics, and information technology—sectors that are economically salient in the state (Texas Higher Education Coordinating Board, 2023b).

2.3 The Texas Core Curriculum

To facilitate transfer from community colleges to public four-year universities and to standardize lower-division academic coursework across institutions, the state has implemented the Texas Common Course Numbering System (TCCNS) and the Texas Core Curriculum (TCC). Established under state statute and overseen by the Texas Higher Education Coordinating Board, TCC requires that all public institutions of higher education —both community colleges and four-year universities — offer and accept a common set of lower-division general education courses (Texas Higher Education Coordinating Board, 2023a). The Core Curriculum consists of 42 semester credit hours across foundational component areas.[1] Institutions have discretion over specific course offerings within each component area, but approved courses must meet statewide learning objectives defined by the Coordinating Board. Completion of the 42-credit core at a community college guarantees transferability of those credits to any public four-year institution in Texas. This creates a common academic baseline across institutions and disciplines, particularly for transfer-oriented students. TCCNS is a system that standardizes course numbering across institutions. All public community colleges in Texas adopt the TCCNS and lower-division academic courses in four-year colleges, especially those included in the TCC, are mapped to TCCNS course numbering.

The existence of a standardized but locally implemented lower-division academic curriculum implies that students at different institutions enroll in courses that are formally comparable in title, credit value, and learning objectives, even though instructional content and materials are determined locally. This combination of statewide standardization and local discretion plays a central role in our setting and motivates our empirical analysis.

3 Data

Our analysis combines information from several sources: the text of over a million course syllabi, detailed records for all students who took these courses, and data on job ads. We describe data source below.

[1] These include communication, mathematics, life and physical sciences, language/philosophy/culture, creative arts, American history, government/political science, and social and behavioral sciences.

3.1 Course Syllabi

Texas's House Bill 2504 requires all public colleges and universities to make course syllabi available to the public and to maintain these records for two years.[2] We compiled a database containing the text of 605,913 syllabi from community college courses by scraping each institution's website. Our data include all current syllabi for 40 of the 50 public community colleges in Texas, and past syllabi (from as early as 2009) for 20 of them.[3]

To compare community college syllabi with those from four-year institutions, we additionally collected around 400,000 syllabi from seven four-year institutions. This sample includes Stephen F Austin State University, Sam Houston State University, Texas A&M, University of Houston - Clear Lake, UT Austin, UT Dallas, and West Texas A&M.

To link courses across institutions, we assign each course the corresponding code in the Texas Common Course Numbering System (TCCNS). All Texas public community colleges use the TCCNS for course numbering in both academic and vocational courses. We assign lower-division four-year college courses a TCCNS code using a crosswalk maintained by the Texas Higher Education Coordinating Board (THECB).[4] Our syllabi data is described in panel (a) of Table 1.

3.2 Students

Information on community college students comes from administrative records of the Education Research Center (ERC). The ERC allows researchers to link detailed data from multiple state agencies covering all students at public K-12 institutions, students at higher-education institutions, and individuals in the labor market in the state of Texas.

Our main source of information is the records of the Texas Higher Education Coordinating Board (THECB), covering all students enrolled in undergraduate and graduate degree programs at public two-year, four-year, and health-related institutions. We extract information on college enrollment and major, transcript-level performance, and degree attainment, available since the 2011–2012 academic year.[5] Students in community colleges are classified as either academic or vocational in the administrative record. Using course transcripts, we link each student to the syllabi of the courses they took while in college. The courses are also classified as either academic or vocational in the

[2] Prior work examining college syllabi (e.g. Nowogrodzki, 2016; Biasi and Ma, 2022; Chau et al., 2023) uses data from the Open Syllabus Project (OSP). OSP has poor coverage on community college syllabi, making it unsuitable for our study.

[3] 10 community colleges had syllabi repositories that were not scrapable with our automated program.

[4] The crosswalk can be accessed through the TCCNS website.

[5] While information on enrollment and degree attainment are available since the 1990s, transcript data, which is crucial to our analysis, is available since the 2011-2012 academic year.

Table 1: Summary Statistics of Syllabi, Students, and Instructors

	Community Colleges		Four-year
	Academic	Vocational	institutions
	(1)	(2)	(3)
Panel (a): Syllabi			
# syllabi	357,686	92,600	11,532
# courses	49,646	49,000	
Average # sections	15.68	5.87	
Average # instructors	6.84	3.25	
Panel (b): Students			
High school exam rank	61.264	56.077	68.504
	(17.670)	(17.832)	(15.288)
Share economically disadvantaged	0.445	0.499	0.328
	(0.497)	(0.500)	(0.470)
High school-county characteristics			
Avg exam rank of high school	60.558	60.428	71.972
	(3.627)	(4.017)	(3.774)
Share economically disadvantaged in high school	0.445	0.422	0.319
	(0.497)	(0.188)	(0.176)
Median income in home county ($)	68,275	69,606	72,969
	(18,239)	(20,426)	(18,883)
Outcomes			
Graduates with Associate Degree	0.186	0.225	
	(0.389)	(0.418)	
Earnings after Associate degree ($)	10,178	10,534	
	(5,347)	(5,605)	
Transfers to four-year	0.263	0.103	
	(0.440)	(0.303)	
Graduates with Bachelor's Degree	0.158	0.061	0.548
	(0.365)	(0.240)	(0.498)
Earnings after Bachelor's degree ($)	12,339	12,463	13,259
	(6,109)	(6,186)	(6,908)
# students	134,223	27,834	801,220
Panel (c): Faculty			
Share female	52.990	46.980	44.480
Share part-time	38.390	29.330	44.170
Share taught in CC			8.590
Share taught in four-year university	27.088	6.315	
Share with BA	5.400	23.940	
Share with MA	70.250	35.040	
Share with PhD	20.730	6.400	
Tenure at Institution	8.494	7.969	5.713
	(6.260)	(6.011)	(5.216)
Salary ($)	38,251	43,781	24,716
	(30,058)	(28,546)	(22,770)

Notes: Summary statistics of students and faculty in the analysis sample. Standard deviations reported in parentheses when relevant.

administrative record. THECB data also contain information on college faculty, including their education, work history, earnings, and courses taught, which we link to student records using transcripts. Course transcripts allow us to observe whether students ever took remedial classes. The administrative records classify students as academically disadvantaged if they took at least one remedial class.

To construct a measure of college students' baseline ability, we link THECB data to records from the Texas Education Agency (TEA), covering all K-12 students in Texas. We use information on mandatory standardized high school test scores in mathematics, reading, and writing.[6] We measure students' baseline ability as their percentile ranking within their cohort. We calculate students' percentile ranking in math, reading, and writing, as well as their average ranking across math and reading. TEA data also provide information on students' eligibility for free or reduced-price lunch, which we use as a measure of family economic disadvantage.

Lastly, we link students' records to information from the Texas Workforce Commission, including quarterly earnings (from unemployment insurance records) for all employed individuals in Texas. We measure earnings in 2024 U.S. dollars.

Panel (b) of Table 1 describes our sample of students at community colleges, separately for students in academic and vocational programs, and compares them to students at four-year institutions. Vocational community college students are more likely to be economically disadvantaged (50%, compared to 45% for academic students and 33% for four-year students) and have lower high school test scores. They also tend to come from lower-performing and more disadvantaged high schools and from counties with lower income. Only 19% of academic students and 23% of vocational students earn an Associate degree as their highest completed degree. About 26% of academic students transfer to a four-year institution and 16% eventually graduate with a bachelor's degree. These figures are 10% and 6%, respectively, for vocational students. Conditional on attaining a BA, community college students earn around $12,000 per quarter one to nine years after graduation, slightly lower than earnings of four-year students and about $2,000 more than community college students who only attain an Associate degree. Our analysis sample contains 134,223 students in academic programs (including 121,276 for whom we have earnings data) and 27,834 students in vocational programs with earnings data.[7]

[6]SAT and ACT test scores are available only for a subset of community college students. High school test scores include the State of Texas Assessments of Academic Readiness (STAAR) for the years 2012-2016; the Texas Assessment of Knowledge and Skills (TAKS) for the years 2003-2015; and the Texas Assessment of Academic Skills (TAAS) for the years 1994-2007. In addition to mathematics, reading, and writing, TAKS and STAAR also have separate modules for science and social sciences; we exclude these from our analysis because they are unavailable for earlier years.

[7]For students in vocational programs, having observed earnings is a requirement for being in the analysis sample

Panel (c) of Table 1 describes the sample of faculty at community colleges and four-year institutions. Most faculty work full-time. Among community college instructors teaching academic courses, 27% previously taught at a four-year college; this share is lower at 6% for instructors teaching vocational courses. Most community college faculty have a Master's degree; only 21% of academic-course faculty and 6% of vocational-course faculty have a Ph.D.

3.2.1 Labor Market Skills

We measure labor market demand for specific workforce skills using information from a database of job postings maintained by Lightcast.[8] For each posting, the database contains information on the year and geographic location, together with a list of skills required for the job drawn from a taxonomy of over 30,000 skills maintained by the company. These skills are further aggregated into 29 clusters. Appendix Figure A1 lists these clusters, together with their frequency across Texas job ads in 2010, 2015, and 2022.

4 Measuring Course Content

We want to understand whether the content of community college courses matches the two stated goals of these institutions: Facilitating transfer to a four-year institution and preparing students for the labor market. To this end, we use the text of course syllabi to construct two measures of course content: (i) *academic relevance*, defined as the similarity of a community college course to four-year courses in the same subject, and (ii) *labor-market relevance*, defined as the extent to which a course covers skills that are in high demand in the labor market. Both measures are constructed using natural language processing techniques applied to the text of syllabi. We describe each measure below.

4.1 Academic Relevance

To measure whether community colleges adequately prepare students for a transfer to a four-year institution, we develop a measure that captures the content similarity between community college courses and their counterpart at four-year colleges. The intuition is that if a community college course is similar in content to a four-year course, then it is more likely to provide students with the knowledge and skills needed to succeed after transfer.

To implement this idea, we create pairs of community college and four-year courses that are equivalent based on the Texas Common Course Numbering System (TCCNS). For example, an in-

[8] since earnings are the primary outcome variable for these students.

troductory psychology course at a community college and a four-year university are both assigned the course number PSYC 2301 under TCCNS, because THECB assesses them as covering similar content and learning objectives. We use the TCCNS to identify these pairs and compare syllabi to measure their content similarity. This implies that we focus on courses with a TCCNS number and exist in both community colleges and four-year universities. In the community college sample, these courses are almost exclusively labeled as academic. In the four-year sample, they are almost exclusively lower-division courses, which are courses designed for first and second year students.

Our measure of similarity is the cosine similarity of word vectors derived from the syllabi. Following Biasi and Ma (2026), we create a vector for each syllabus. We start with a term dictionary, constructed as the set of all words, bigrams, an trigrams ever appearing as keywords in all academic publications in the OpenAlex database (see Appendix B for details). We construct each word vector as having length equal to the total number of n-grams in the dictionary; elements are equal one if the document contains a given n-gram and zero otherwise. To account for differences in use across n-gram, we further implement a Text Frequency–Backward Inverse Document Frequency (TF-BIDF) weighting scheme (Biasi and Ma, 2022). This procedure multiplies each element by its frequency within the document (TF) and divides it by its frequency across all documents published up to its date (BIDF). Using these weighted word vectors, we finally calculate the cosine similarity between pairs of syllabi. This metric ranges from -1 to 1, with higher values indicating greater similarity in content.

We calculate each community college course's academic relevance as the average cosine similarity to all four-year syllabi with the same TCCNS number, across all institutions. Additionally, we construct the average across courses taught at different types of institutions (UT Austin, R1, and non-R1). For comparison, we also calculate the average cosine similarity between four-year syllabi with the same TCCNS number across all four-year institutions. This allows us to benchmark the academic relevance of community college courses relative to the variation in content across four-year institutions.

4.2 Labor Market Relevance

We also want to measure whether community college courses adequately prepare students for the labor market. To do so, we first use job ads data from Lightcast to construct a measure of labor market relevance of each of 29 skill cluster families in the platform's taxonomy (Appendix Figure

A1).[9] We then use this measure to assess the labor market relevance of community college courses. Our approach can be summarized as follows:

1. For each county j in Texas and year t, we calculate the share of job postings that mention each skill cluster k in the Lightcast taxonomy. This gives us a measure of the demand for each skill cluster in the local labor market. We call this measure ω_{jkt}. We further normalize it so that it sums to one across skill clusters for each county and year, which gives us $\omega^*_{jkt} \equiv \frac{\omega_{jkt}}{\sum_k \omega_{jkt}}$.

2. For each community college course s, we count the frequency of mentions of skills in each skill cluster k in the course syllabus. This gives us a measure of the skill content of the course. We call this measure ϕ_{sk}.

3. We define the labor market relevance of course s as the weighted sum of the skill content of the course, where the weights are given by the demand for each skill cluster in the local labor market. Formally, we define it as

$$\Phi_s = \sum_k \phi_{sk} \omega^*_{j(s)kt(s)},$$

where $j(s)$ is the county where the community college offering course s is located and $s(t)$ is the year of instruction. This measure captures the extent to which community college courses cover skills that are in high demand in the local labor market in that same year.

This aggregate measure can also be decomposed into skill-specific components. Rather than summing across all clusters, we can isolate the density of a single cluster k, ϕ_{sk}. This skill-specific version allows us to examine which particular types of skills—such as IT, healthcare, or business—drive the effects of labor-market relevance on student outcomes, and to test for heterogeneity across skill clusters.

5 Differences in Course Content Across Courses

Applying these measures to data from Texas course syllabi, we begin by studying how the academic and labor-market relevance of community college courses vary across courses. We first describe the distribution of our measures across community college courses, and compare them to four-year college courses. We then examine how these measures vary across fields and instructors, and

[9] Lightcast is a labor-market analytics firm that aggregates and processes data from online job postings and other sources to produce datasets on occupations and skills demand. The company was formed from the merger of EMSI and Burning Glass Technologies and was previously known as Emsi Burning Glass.

how they correlate with student characteristics. While being informative in itself, this descriptive evidence serves as a validation of our measures by providing an assessment of their validity and reliability, and also provides the foundation of our causal analysis on the effects of college courses on a range of student outcomes.

5.1 Academic Relevance (AR)

Community college courses are more similar to courses at four-year institutions than four-year courses are to one another. On average, community college courses have a cosine similarity of 0.181 to courses with the same TCCNS number across the seven four-year universities in our sample (Figure 1, panel (a)). By comparison, the average cosine similarity among four-year courses is 0.168 (s.d. 0.077; Figure 1, panel (b)). This pattern holds when comparing syllabi to different subsets of four-year universities. The average similarity of community college courses falls to 0.138 when the comparison is limited to the state flagship, UT Austin, and rises to 0.188 when it is limited to the non-R1 universities in the sample, indicating that community college courses are more similar to those at non-R1 universities than to those at the flagship. Within the four-year sector, courses at UT Dallas and Texas A&M — the other two R1 universities in our sample — have an average similarity of 0.165 to corresponding courses at UT Austin. Courses at non-R1 universities are less similar to those at UT Austin (0.125) but more similar to those at other non-R1 universities (0.172).

Despite the overall similarity to four-year courses, community college courses differ significantly in their AR. Moving a community college course from the 25th percentile to the 75th percentile of the distribution — a 0.1 difference in AR — is equivalent to replacing 16% of the content of that syllabus with content from the four-year syllabus (Figure 1, panel (a)).[10]

The overall similarity may also mask variation within courses across target universities within types. If there is enough variation in course content across courses at four-year universities, a community college course may be very similar to courses from a specific four-year institution but very different from all the others, resulting in an overall low average similarity. To address this issue, we compute — for each course — the *maximum* similarity to any four-year course with the same TCCNS number. The average maximum similarity of community college courses to four-year courses rises to 0.23, higher than the average similarity (Appendix Figure A3). Nevertheless, this version of AR is still higher for community colleges than for four-year institutions (0.168).

AR differs across fields as well. Community college courses in the Social Sciences have the

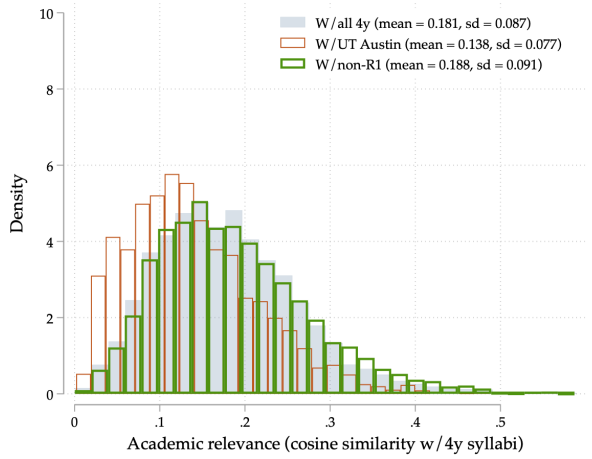
[10] To see this, notice that if the initial content of the syllabus has a similarity of 0.08 with the average four-year syllabus and we want it to have a similarity of 0.23, the replaced share x must satisfy $0.23 = (1-x)0.11 + x$, which implies $x = 0.163$.

highest average AR at 0.204, followed by Science courses at 0.196 and Humanities and other courses at 0.161 (Appendix Figure A2). When considering similarity to UT Austin courses only, AR is the lowest for Sciences courses; when considering non-R1, it is the highest for Sciences. By comparison, four-year courses in the Social Sciences have an average AR of 0.173, Science courses are at 0.179, and Humanities and other courses at 0.157.

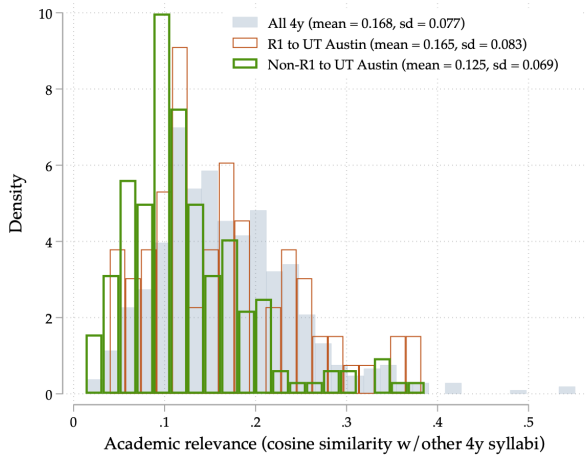
Overall, these patterns suggest that community college courses are broadly aligned with four-year curricula, especially outside the flagship campus.

Figure 1: Distribution of Academic Relevance, by Institution Type

(a) Community colleges



(b) Four-year institutions



Notes: Histogram of AR across courses at community colleges and at four-year institutions in our sample, overall and by type of comparison institution. AR is defined as a course syllabus's average cosine similarity with 4-year courses with the same code within the Texas Core Curriculum.

Decomposing the variance in academic relevance. To understand the sources of variation in AR, we perform a Shapley-Owen decomposition of the variance in AR into the following factors: school year, field, course, and instructor. This procedure quantifies the portion of total variation that can be attributed to each factors by estimating a series of regressions of AR on various combinations of the sets of fixed effects pertaining to each factors and then measuring the average decline in adjusted R^2 of these regressions when each factor's fixed effects are removed.[11]

[11]The decomposition involves three steps: (i) estimate OLS models of proximity on all possible combinations of factors and record their adjusted R^2; (ii) re-estimate these models excluding one factor at a time and record the decrease in adjusted R^2; (iii) calculate the average decline in adjusted R^2 for each factor across all possible combinations, defining this average as the partial-R^2 attributable to each factor. For factor j, the partial-R^2 is calculated as $R^2_j = \sum_{T \subseteq V \setminus \{j\}} \frac{|T|!(K-|T|-1)!}{K!}[R^2(T \cup \{j\}) - R^2(T)]$, where $R^2(S)$ is the adjusted R^2 of a regression of proximity on the set of factors S, V is the full set of factors, $|T|$ denotes the number of factors in subset T, and $K \equiv |V| = 5$ is the total number of factors. Using adjusted R^2 ensures comparability across factors with different numbers of categories.

Table 2: Shapley-Owen Decomposition of Variance in Academic and Labor Market Relevance

	Academic Relevance		Labor Market Relevance			
	Academic courses		Academic courses		Vocational courses	
	Partial R^2	% explained	Partial R^2	% explained	Partial R^2	% explained
	(1)	(2)	(3)	(4)	(5)	(6)
Year	0.012	1.6	0.081	9.99	0.016	3.02
Field	0.352	47.52	0.083	10.21	0.224	42.97
Course	0.328	44.27	0.075	9.24	0.138	26.42
Institution	0.041	5.47	0.517	63.62	0.112	21.42
Instructors	0.008	1.15	0.056	6.95	0.032	6.18
Total	0.740	100	0.812	100	0.522	100

Notes: This table presents the results of a Shapley Owen decomposition of the variance in academic relevance and labor market relevance for academic and vocational community college courses. The decomposition attributes the variance to different factors, including school year, field, course, institution, and instructors. The *partial* R^2 indicates the proportion of total explained variance attributable to each factor, while the *percent explained* column shows the percentage of total explained variance attributed to each factor.

This exercise confirms that most of the differences in courses' AR occur across fields. About 35% of the total variation in AR and 48% of the explained variation can be attributed to differences across field. Differences across courses account for 44% of explained variation, indicating a significant degree of persistence in this measure within courses over time. By contrast, differences across institutions account for only 5.5% of total variation, while differences across instructors account for just over 1%. This limited role of instructors in shaping this dimension of course content contrasts with previous work examining the determinants of the presence of frontier knowledge in four-year courses, finding instructors to be one of the main drivers (Biasi and Ma, 2022).

AR and student, instructor, and course characteristics. To better understand what aspects of courses are associated with higher AR, we examine its relationship with student and instructor characteristics by regressing this measure on school-by-field-by-year fixed effects and the following variables: students' average rank in the cohort distribution of a composite high school exam score, a measure of ability; the share of students who are academically disadvantaged; the share of students who are economically disadvantaged; indicators for the instructor's highest education degree (bachelor's, master's, or Ph.D.); an indicator for whether the instructor ever taught at a four-year university within Texas; an indicator for the instructor being employed part-time; and the instructor's years of experience in the community college. This exercise shows that AR is higher for courses enrolling fewer economically disadvantaged students. It is also higher for courses taught by in-

structors with a Ph.D, by those with more experience teaching at the college, and by those working full-time. These patterns are consistent with the idea that the content of community colleges may be shaped by the characteristics of the students and the experience and expertise of instructors.

Table 3: Course Content, Student Characteristics, and Faculty Characteristics

	Academic relevance (sd)		Labor market relevance (sd)			
	Academic courses		Academic courses		Vocational courses	
	Levels	Δ	Levels	Δ	Levels	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
Students						
High school rank	0.000	-0.005	-0.017*	-0.001	-0.000	-0.002
	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)
Share acad. disadv.	-0.028	0.068*	-0.048*	0.016	0.026*	-0.012
	(0.024)	(0.032)	(0.023)	(0.029)	(0.012)	(0.019)
Share econ. disadv.	-0.104*	-0.053	0.016	-0.073*	-0.049*	0.004
	(0.026)	(0.036)	(0.025)	(0.031)	(0.013)	(0.020)
Faculty						
Bachelor's	-0.047	-0.085	-0.134*	0.015	0.000	-0.007
	(0.039)	(0.052)	(0.038)	(0.047)	(0.009)	(0.014)
Master's	0.025	-0.023	-0.126*	-0.056	0.024*	0.007
	(0.030)	(0.040)	(0.029)	(0.036)	(0.010)	(0.014)
Ph.D	0.053*	-0.032	-0.080*	-0.045	0.039*	0.056*
	(0.031)	(0.041)	(0.030)	(0.037)	(0.015)	(0.022)
Taught at 4-year	-0.004	-0.009	0.012	-0.004	0.002	-0.017
	(0.009)	(0.012)	(0.009)	(0.011)	(0.012)	(0.017)
Years of experience in CC	0.002*	0.000	-0.000	0.002	-0.002*	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Part-time	-0.021*	-0.014	0.046*	0.003	0.067*	-0.020*
	(0.010)	(0.013)	(0.010)	(0.012)	(0.008)	(0.011)
N	36,562	34,538	42,220	40,414	64,675	50,474
R^2	0.398	0.009	0.321	0.007	0.416	0.006
F-stat	11.590	1.310	15.900	2.100	36.020	1.880

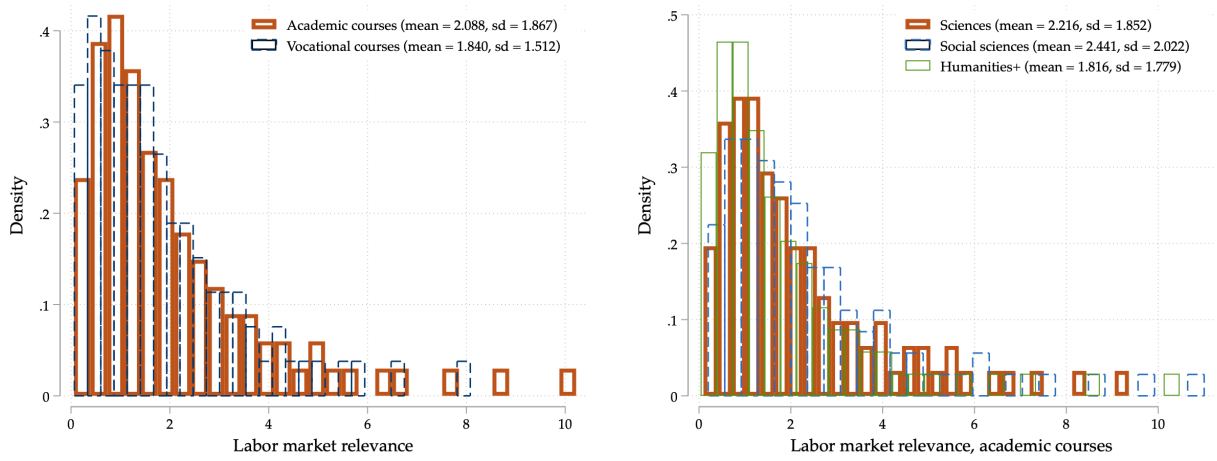
Notes: OLS estimates of a set of student and faculty characteristics on academic and labor market relevance of community college courses. One observation is a course. Academic relevance is defined as the average cosine similarity with courses at four-year institutions with the same TCCNR. Labor market relevance is defined as in the text. Both measures are standardized to have mean 0 and variance 1. Columns 1-4 focus on academic courses, columns 5-6 on vocational courses. In columns 1, 3, and 5 the dependent variable is in levels; in columns 2, 4, and 6 it is in changes from the previous year. All specifications control for institution-by-field-by-year fixed effects. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.1.1 Labor Market Relevance (LMR)

Next, we examine the distribution of LMR across community college courses, distinguishing between academic and vocational courses. The average LMR of academic community college courses

is 2.09, with a standard deviation of 1.87. This figure is higher than the average LMR of vocational courses, which is 1.84 (s.d. 1.51). As for AR, significant differences exist across fields: among academic courses, LMR is highest in the Social Sciences at in 2.44, followed by Science courses at 2.22 and Humanities and Others at 1.82.

Figure 2: Distribution of Labor Market Relevance, by Course Type and Field



Notes: Histogram of LMR across courses at community colleges, as defined in the text.

Among academic courses, most of the variation in LMR is differences across colleges: A Shapley-Owen decomposition indicates that the overwhelming majority of the explained variation in this attribute, 64%, is explained by institutions. By comparison, differences across field account for 10% of the explained variation, courses for 9%, and instructors for 7% (Table 2, columns 2-3). Among vocational courses, most of the variation (43%) can be attributed to field differences, followed by courses (26%), institutions (21%), and instructors (6%).

Lastly, we examine the relationship between LMR and student and instructor characteristics. For academic courses, LMR is lower for courses enrolling more academically disadvantaged students. It is also lower for courses taught by instructors with bachelor's, master's, or Ph.D degrees and by those working full-time. For vocational courses, LMR is higher for courses enrolling fewer economically disadvantaged students and more academically disadvantaged students, and for courses taught by instructors with a master's or Ph.D and working part-time. These patterns suggest that the LMR of community college courses may be shaped by the characteristics of the students they serve and the experience and expertise of their instructors, in ways that relate to the ultimate purpose of the course: preparing students for a four-year college education or for the labor market.

6 Estimating The Effect of Course Content on Student Outcomes: Research Design

The syllabi data show a significant amount of variation in our two measures of community college course content. Do these differences matter for student outcomes? Studying this question requires solving a fundamental identification issue: students choose which courses to take, in ways that could potentially relate to their future outcomes. In this section, we outline our baseline model and then propose two alternative strategies to tackle this challenge.

6.1 Baseline Model

Consider student i, enrolled at community college j. Let C_{ij} be the collection of all courses the student takes while in college, and denote each course as $c_{njt} \in C_{ij}$, where n is the course numbering and t is the year of instruction. Each course has some measure of content $m(c_{njt})$ (either AR or LMR). Let $\bar{m}_{ij}$ denote the average measure of all courses in the set C_{ij}: $\bar{m}_{ij} \equiv \sum_{c_{njt} \in C_{ij}} m(c_{njt})/|C_{ij}|$, where $|C_{ij}|$ denotes the total number of courses the student takes. Denoting y_j as an outcome measure (for example earnings), the model we want to estimate is:

$$y_i = \beta \bar{m}_{ij} + \delta X_{ij} + \epsilon_{ij} \tag{1}$$

where X_{ij} is a vector of controls specific to the student (for example gender, race or ethnicity, college fixed effects, or cohort fixed effects). Students' selection of courses implies that $\bar{m}_{ij}$ could be correlated with ϵ_{ij}, leading to bias in OLS estimates of β.

6.2 Main Identification Strategy: Last-minute updates to course syllabi

Our first identification strategy leverages last-minute updates to the content of a course, unobservable to the students when they choose which courses to take. Texas HB 2504 requires course instructors to post their updated course syllabus within one week from the start of the course, but not before (see Appendix C for details). This implies that, at the time of course registration, students can observe the characteristics of the instructor and the syllabus from the previous term, but not the updated one; i.e., they observe $m(c_{njt-1})$ but not $m(c_{njt})$. This implies that, conditional on $m(c_{njt-1})$ and instructor characteristics, $m(c_{njt})$ is as good as randomly assigned. We implement this identification strategy by including $\bar{m}_{ij}^{lag} \equiv \sum_{c_{njt} \in C_{ij}} m(c_{njt-1})/|C_{ij}|$, i.e., the average content measure of student i's courses measured in the previous term, in equation (1)

$$y_i = \beta \bar{m}_{ij} + \gamma \bar{m}_{ij}^{lag} + \delta X_{ij} + \epsilon_{ij}. \tag{2}$$

To account for the fact that, when choosing whether to take a specific course, students take into account all courses available to them, we control for college-by-major-by-starting cohort fixed effects; the course choice set should be the same within these cells.

There are two identification assumptions underlying this strategy. The first is that students cannot sort across courses with different content updates based on ability. This assumption is supported by the institutional framework outlined by Texas HB 2504. Additionally, in Table 3 (columns 2, 4, and 6) we show that one-period changes in course content measures cannot be predicted by student and instructor characteristics (the F-statistics are all lower than 3). The second is that changes in course content are not systematically related to other course changes that may affect outcomes—for example changes in the course instructor. To meet this assumption, all specification control for time-varying instructor characteristics (education degree, experience at the community college, an indicator for having taught at a four-year institution, and part-time status).

6.3 Alternative Identification Strategy: Proximity to community colleges with different content

Our second identification strategy combines variation in course content across community colleges in different areas of Texas with the established fact that the vast majority of students enroll in community colleges that are the closest to where they live (e.g., Mountjoy, 2022). Defining $k(i)$ as the college closest to i, we construct $\bar{m}_i^{closest} \equiv \frac{\sum_{c \in C_{k(i)}} m(c)}{|C_{k(i)}|}$, the average measure of content of all courses at k. We then use this variable as an instrument for $\bar{m}_{ij}$ in equation (1) and estimate the equation's parameters via 2SLS.

For this strategy to yield consistent estimates of β, a fundamental exclusion restriction must be satisfied: $\bar{m}_i^{closest}$ must not be correlated with ϵ_{ij} conditional on X_{ij}. This restriction is violated if college instructors adjust content based on the characteristics of the pool of prospective students in the local community: for example, they offer easier (or less academically relevant) material when students are more likely to be academically or economically disadvantaged). To bolster the identification assumption, we control for fixed effects for the locale code of the student's high school; median income in the student's home county; number of full-time equivalent teachers in the student's high school; number of economically disadvantaged students in the student's high school; and the share of minority students in the student's high school.

We consider the exclusion restriction to be plausible when $m(\cdot)$ is AR. However, community colleges may choose to emphasize a specific skill in response to the local labor market demands. As a result, the exclusion restriction may be violated when $m(\cdot)$ is LMR and, especially, when $m(\cdot)$ is specific skill families. For this reason, we prioritize our main identification strategy and use this 2SLS strategy as a robustness check.

7 Course Content and Student Outcomes

We now turn to the central question of the paper: does course content affect the outcomes of community college students? We examine impact of our two dimensions of content in turn — academic relevance (AR) and labor-market relevance (LMR) — on degree attainment, transfer to four-year institutions, academic performance after transfer, and earnings. In section 7.5, we also examine the effects of specific skill clusters. For the effects of LMR and skill clusters, we separately examine the results for students in academic and vocational tracks. Our main estimates are based on the lagged-content specification in equation (2), which exploits last-minute updates to syllabi that are unobserved by students at enrollment.

7.1 Academic Relevance: Degree Attainment, Transfer, and Earnings

We begin with the effects of AR on students enrolled in academic programs. Table 4 reports OLS estimates of equation (2), where the independent variable is a student's average AR across all courses taken at the community college, standardized to have mean zero and variance one across all students. In all specifications, we control for lagged academic relevance; college-by-major-by-cohort fixed effects; share of the student's instructors having a Ph.D. and having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, underrepresented minorities, student-staff ratio, county median income, and locale. For the specification with log earnings as dependent variable, we additionally control for a quadratic of potential experience, defined as numbers of years since last observed college enrollment. Columns 1–3 report effects on binary indicators for associate degree attainment, transfer to a four-year institution, and bachelor's degree attainment, respectively. Column 4 reports the effect on GPA at the four-year institution, estimated on the subsample of students who transfer. Column 5 reports the effect on log quarterly earnings, pooled over the 1–9 years following graduation.

Associate degree and bachelor's degree attainment. Higher AR significantly raises the probability of earning a degree at both levels. A one-sd increase in a student's average AR raises the

Table 4: Effects of Academic Relevance on Student Outcomes

	Associate degree	Transfer to 4-year	Bachelor's degree	GPA at 4-year	Log earnings
	(1)	(2)	(3)	(4)	(5)
Academic relevance (sd)	0.011***	0.009***	0.015***	0.016***	0.015***
	(0.002)	(0.003)	(0.002)	(0.006)	(0.005)
Mean dep. var.	0.186	0.263	0.158	2.985	
R^2	0.393	0.338	0.246	0.173	0.254
N	126,971	126,971	126,971	33,189	321,840
N clusters					31,907

Notes: OLS estimates of β in equation (2). The dependent variable is an indicator for obtaining an associate degree (column 1), for transferring to a four-year institution (column 2), for obtaining a bachelor's degree (column 3), GPA at the four-year institution conditional on transfer (column 4), and the natural logarithm of quarterly earnings 1–9 years post-graduation (column 5). The independent variable is the measure of average academic relevance across a student's courses, as defined in Section 4, standardized to have mean zero and variance one. In columns 1–4, one observation is a student. In column 5, one observation is a student-quarter, and standard errors are clustered at the student level; robust standard errors are used in columns 1–4. All specifications control for lagged academic relevance; for college-by-major-by-cohort fixed effects; for share of student's instructors having a Ph.D. and having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, under-represented minorities, student-staff ratio, county median income, and locale. For the specification with log earnings as dependent variable, we additionally control for a quadratic of potential experience, defined as years since last observed college enrollment. The sample is restricted to students in academic programs. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

probability of completing an associate degree by 1.1 pp (Table 4, column 1, significant at the 1% level). Relative to a baseline completion rate of 18.6%, this represents a 5.9% increase. The effect on bachelor's degree attainment is larger in relative terms: 1.5 pp, also significant at the 1% level, or a 9.5% increase relative to the mean of 15.8%, .

These effects operate through two complementary channels. AR raises both the propensity to transfer and the ability to complete a degree upon transfer. Even conditional on eventual transfer, students whose community college courses were more academically aligned are more likely to complete a four-year degree, consistent with better academic preparation reducing dropout and time-to-completion at the receiving institution.

Transfer to four-year institutions. AR also significantly raises the probability of transferring to a four-year institution: the effect of a one-sd increase in AR is 0.9 pp, significant at the 1% level, corresponding to a 3.4% increase relative to the baseline transfer rate of 26.3% (Table 4, column 2, significant at 1%). Transfer is a necessary precondition for bachelor's degree attainment, and this result partly explains the positive effect on BA completion. At the same time, the magnitude of the

BA effect (1.5 pp, Table 4, column 3) exceeds what the transfer effect (0.9 pp) alone would predict if all additional transferees graduated at the baseline rate, suggesting that AR also raises completion rates among students who would have transferred regardless.

Academic performance after transfer. Among students who do transfer, higher AR leads to small improvements in academic performance at the receiving institution. A one-sd increase in academic relevance raises post-transfer GPA by 0.016 points, against a baseline GPA of 2.985 (Table 4, column 4, significant at 1%). This result is identified in the subsample of 33,189 students with observed four-year enrollment. The finding suggests that the academic preparation imparted by syllabus alignment extends beyond the community college gates: students arrive better equipped for the pace and rigor of four-year coursework.

Earnings. AR also has a direct effect on earnings. A one-sd increase in this measure raises earnings by approximately 1.5% (Table 4, column 5, significant at 1%). This estimate is obtained from a sample of 321,840 student-quarter observations covering up to nine years of post-graduation labor market experience. The positive earnings effect is consistent with the educational attainment results: higher AR raises degree completion and transfer rates, and the credentials and skills acquired through more rigorous preparation translate into better labor market outcomes. Notably, the earnings effect is identified through variation in within-course content changes, providing causal evidence that academic preparation — not unobserved student characteristics — drives the improvement.

7.2 Alternative Strategy: Distance-Based Instrument

As a second identification strategy, we instrument for each student's average academic relevance using the average academic relevance of all courses offered at the community college geographically closest to the student's home, as described in Section 6. The instrument exploits the strong tendency of students to enroll near where they live (Mountjoy, 2022): proximity to a high-relevance college raises the probability of exposure to more academically aligned instruction, independently of the student's own choices or characteristics. Because the exclusion restriction requires that local college content does not reflect local student composition, we include a rich set of controls for local area characteristics as described in Section 6 and above. The first-stage F-statistic of 346 confirms that proximity to a high-AR college is a powerful predictor of students' realized AR (Table 5).

The 2SLS estimates are broadly consistent but larger in magnitude with those from the lagged-content specification in Table 4. A one-sd increase in academic relevance raises the probability of

Table 5: Effects of Academic Relevance on Student Outcomes

	Associate degree	Transfer to 4-year	Bachelor's degree	Log earnings
	(1)	(2)	(3)	(4)
Academic relevance (sd)	0.062***	0.010	0.034***	0.034
	(0.007)	(0.007)	(0.005)	(0.029)
R^2	0.406	0.337	0.227	0.252
N	81,550	81,550	81,550	99,341
N clusters				32,538
First-stage F-stat		345.960		

Notes: 2SLS estimates of β in equation (2). The dependent variable is an indicator for obtaining an associate degree (column 1), for transferring to a four-year institution (column 2), for obtaining a bachelor's degree (column 3), and the natural logarithm of quarterly earnings 1–9 years post-graduation (column 4). The independent variable is the measure of average academic relevance across a student's courses, as defined in Section 4 and standardized to have mean 0 and variance 1. This variable is instrumented by the average relevance at the community college that is closest to each student; the reported F-statistic refers to this excluded instrument. In columns 1–3, one observation is a student. In column 5, one observation is a student-quarter, and standard errors are clustered at the student level; robust standard errors are used in columns 1–4. All specifications control for major-by-cohort fixed effects; for indicators for instructors having a Ph.D. and for having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, under-represented minorities, student-staff ratio, county median income, and locale. For the specification with log earnings as dependent variable, we additionally control for a quadratic of potential experience. The sample is restricted to students in academic programs. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

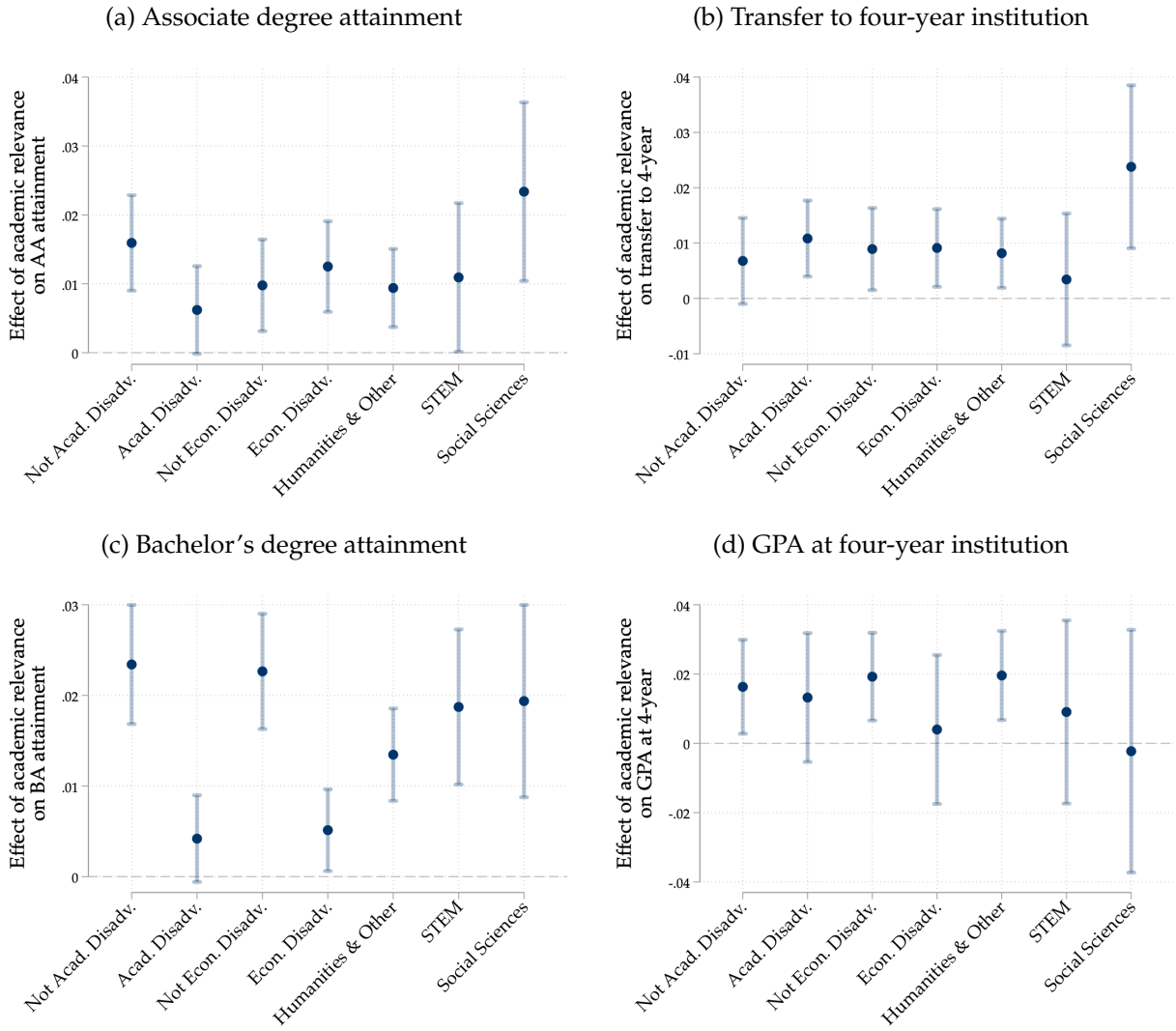
earning an associate degree by 6.2 pp (significant at the 1% level) and of completing a bachelor's degree by 3.4 pp (significant at the 1% level). The effects on four-year transfer and on earnings are positive and large, but imprecisely estimated (1.0 pp and 3.4%, respectively). Combined with the estimates in the previous subsection, these findings confirm that AR of community college courses has positive effects on average on students in academic programs.

7.3 Heterogeneity in the Effects of Academic Relevance

The average effects reported in Table 4 may conceal important heterogeneity across student groups and fields of study. To explore them, we re-estimate equation (2) separately by student's economic and academic disadvantage and by field.

Figure 3 presents coefficient plots for degree attainment, transfer, and post-transfer GPA, and Figure 4 presents earnings results (alongside LMR earnings effects), each estimated separately by academic disadvantage, economic disadvantage, and broad field of study (Humanities and Other, STEM, and Social Sciences).

Figure 3: Effects of Academic Relevance on Degree Attainment and Post-Transfer Performance, by Subgroup



Notes: Each panel plots OLS estimates and 95% confidence intervals of the parameter β in equation (2), estimated separately using the indicated outcome as the dependent variable, standardized average academic relevance as the main independent variable, and on the subsamples indicated on the horizontal axis. Subsamples are defined by academic disadvantage status (not academically disadvantaged, academically disadvantaged), economic disadvantage status (not economically disadvantaged, economically disadvantaged), and broad field of study (Humanities & Other, STEM, Social Sciences). All specifications control for lagged academic relevance; college-by-major-by-cohort fixed effects; indicators for instructors having a Ph.D. and for having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, under-represented minorities, student-staff ratio, county median income, and locale. The sample is restricted to students in academic programs. Confidence intervals are constructed using robust standard errors.

Heterogeneity by disadvantage status. The effects of AR on degree attainment are positive across all subgroups, but larger for non-academically-disadvantaged students. A one-sd increase in AR

increases associate degree attainment effect by 1.6 pp for non-academically-disadvantaged students and by 0.6 pp for disadvantaged ones (Figure 3, panel (a)). The pattern even is more pronounced for bachelor's degree attainment: the effect for non-academically-disadvantaged students (2.3 pp) is roughly five times larger than for academically disadvantaged students (0.4 pp, panel (c)). For this outcome, a near-identical pattern holds by economic disadvantage status. Non-economically-disadvantaged students benefit substantially more from higher AR in terms of BA attainment (approximately 2.3 pp) than disadvantaged students (approximately 0.5 pp). Effects on transfer probabilities are more uniformly distributed across academic and economic disadvantage status (Figure 3, panel (b)).

This pattern has a natural interpretation. Bachelor's degree completion requires not only transferring but also persisting through two or more additional years of four-year coursework — a challenge that depends on both academic preparation and a host of non-academic constraints, such as financial stability, family responsibilities, and access to support networks. For academically and economically disadvantaged students, these non-academic barriers may be sufficiently binding that improved course preparation alone cannot shift BA completion rates. For better-prepared and more financially stable students, in contrast, course quality is more likely the margin that matters.

For GPA at the four-year institution, the effect of AR is generally positive but is notably smaller and statistically indistinguishable from zero for economically disadvantaged students (0.005 GPA points, with a confidence interval spanning zero) compared to non-disadvantaged students (approximately 0.020 GPA points; (Figure 3, panel (d)). This is consistent with the BA attainment results: among economically disadvantaged students who do manage to transfer, the benefits of academic alignment translate into smaller improvements in post-transfer academic performance, perhaps because economic constraints continue to create headwinds even after the transfer hurdle is cleared.

The heterogeneity reverses, however, for earnings (Figure 4). The positive earnings effect of academic relevance is if anything *larger* for academically disadvantaged students (2.8%) than for non-disadvantaged students (1.7%), although the two estimates are statistically indistinguishable from each other. The effect is also similar in magnitude for economically disadvantaged (2.3%) and non-disadvantaged students (2.5%). This suggests that even among disadvantaged students who do not complete a four-year degree, more academically rigorous community college preparation generates returns in the labor market, through enhanced human capital if not through a degree premium.

Heterogeneity by field. Across fields, the effects of AR are positive for all three broad groups. For associate degree attainment, the effect is largest in Social Sciences (2.3 pp), consistent with the finding in Section 5 that Social Sciences courses exhibit the highest AR on average and may have more room for within-course variation to matter (Figure 4, panel (a)). Effects in STEM (1.1 pp) and Humanities and Other (1.0 pp) are smaller but precisely estimated. For transfer, field differences are more muted, with Social Sciences again showing the largest effect (2.4 pp) and STEM the smallest (0.3 pp, with a confidence interval that includes zero (Figure 4, panel (b)). For bachelor's degree attainment, estimates are positive and statistically significant across all three fields (Humanities & Other: 1.3 pp; STEM: 1.9 pp; Social Sciences: 2.0 pp), with relatively little dispersion (Figure 4, panel (c)). For GPA at the four-year institution, effects are positive in STEM and Humanities but near zero in Social Sciences, suggesting that the channels through which Social Sciences AR improves outcomes operate primarily through completion and transfer rather than through the academic performance of those who transfer (Figure 4, panel (d)). Earnings effects, shown in panel (a) of Figure 4, are positive across all fields, with the largest point estimate in Social Sciences (5.5%), though the wide confidence interval for that group precludes strong inference.

7.4 Labor-Market Relevance Effects

We now study the effects of LMR on student outcomes. Because academic and vocational programs are designed with fundamentally different goals — transfer preparation versus direct labor-market entry — we estimate effects separately for students enrolled in each type of program, and for different outcomes: academic measures and earnings for students in academic programs, and earnings for students in vocational programs.

7.4.1 Academic Program Students

For students enrolled in academic programs, higher LMR significantly reduces both degree attainment and the probability of transfer. A one-sd increase in average LMR reduces the probability of completing an associate degree by 3.6 pp (Table 6, column 1, significant at the 1% leve). Relative to a baseline of 18.6%, this represents a 19% reduction. The effect on bachelor's degree attainment is even larger: -4.2 pp or a 27% reduction relative to the baseline completion rate of 15.8% (column 2). Transfer to a four-year institution decreases by 1.4 pp, significant at the 5% level (column 3).

These results are consistent with a content mismatch mechanism. Academic program students are, by definition, pursuing credentials designed to facilitate further study. When their courses are oriented toward labor-market skills rather than the analytical and subject-matter preparation

Table 6: Effects of Labor-Market Relevance on Student Outcomes

	Academic programs					Vocational
	Associate degree	Transfer to 4-year	Bachelor's degree	GPA at 4-year	Log earnings	Log earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Labor market relevance (sd)	-0.036***	-0.014**	-0.042***	0.048***	-0.003	-0.002
	(0.006)	(0.007)	(0.004)	(0.015)	(0.015)	(0.021)
Mean dep. var.	0.186	0.263	0.158	2.985		
R^2	0.394	0.335	0.242	0.166	0.260	0.198
N	138,128	138,128	138,128	36,133	319,066	125,209
N clusters					31,319	12,095

Notes: OLS estimates of β in equation (2). The dependent variable is an indicator for obtaining an associate degree (column 1), for transferring to a four-year institution (column 2), for obtaining a bachelor's degree (column 3), GPA at the four-year institution conditional on transfer (column 4), and the natural logarithm of quarterly earnings 1–9 years post-graduation (columns 5–6). The independent variable is the measure of average labor-market relevance across a student's courses, as defined in Section 4. In columns 1–4, one observation is a student. In columns 5–6, one observation is a student-quarter, and standard errors are clustered at the student level; robust standard errors are used in columns 1–4. All specifications control for lagged academic relevance; for college-by-major-by-cohort fixed effects; for indicators for instructors having a Ph.D. and for having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, under-represented minorities, student-staff ratio, county median income, and locale. For the specification with log earnings as dependent variable, we additionally control for a quadratic of potential experience. The sample in columns 1–5 is restricted to students in academic programs; the sample in column 6 is restricted to students in vocational programs. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

required for four-year coursework, they arrive at the transfer threshold — or the associate degree threshold — with a skill set that is misaligned with academic demands. Crucially, our identification strategy captures the causal impact of last-minute content changes that students did not anticipate and could not have compensated for through additional preparation, lending a clean causal interpretation to this mismatch.

A seemingly counterintuitive result emerges for post-transfer GPA: among academic students who do transfer, higher labor-market relevance is associated with *higher* four-year GPA (0.048 points, Table 6, column 4, significant at the 1% level). This positive conditional estimate may reflect selection. When LMR is high, a smaller and more selected subset of students clears the transfer hurdle: only those with particularly strong academic motivation and ability manage to transfer despite receiving less academically oriented preparation. These highly selected transferees then perform relatively well at four-year institutions.

For post-graduation earnings, the effect of LMR on academic-program students is near zero and

not statistically significant (-0.003; Table 6, column 5). The absence of a significant earnings penalty, despite the large negative effects on degree attainment, could reflect two offsetting forces. On the one hand, students with higher LMR who fail to complete four-year degrees miss out on substantial degree premia. On the other hand, those same students may enter the labor market sooner and with more immediately applicable vocational skills, partially offsetting the credential loss. The two effects appear to roughly cancel in the aggregate, leaving the average earnings coefficient close to zero.

7.4.2 Vocational Program Students

For students enrolled in vocational programs, the primary outcome of interest is post-graduation earnings, as these students are oriented primarily toward labor-market entry rather than further study. The average effect of LMR on the log earnings of vocational students is -0.002, statistically indistinguishable from zero (Table 6, column 6). On average, vocational students neither gain nor lose from higher LMR in their courses.

This null average, however, conceals striking heterogeneity across fields, documented in panel (c) of Figure 4. Among vocational students in Humanities and Other fields, a one-sd higher LMR raises earnings by approximately 11%, consistent with the view that in fields where vocational preparation is less common or where employers are less likely to expect it, course content that explicitly targets labor-market skills confers a meaningful advantage. By contrast, a one-sd increase in LMR reduces earnings of students in STEM fields by 22% and of Social Sciences students by 27% log points, though the latter is less precisely estimated.

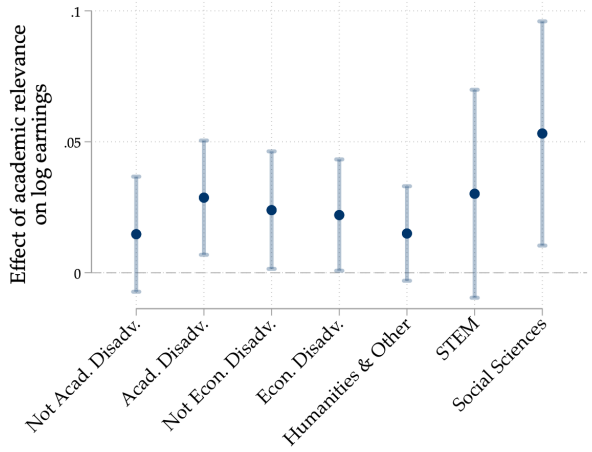
The negative LMR effects for STEM and Social Sciences vocational students are particularly striking. One interpretation is that students in these more specialized fields benefit from specific skills that are narrowly relevant to their target job categories. As a result, having a higher overall LMR spreads the focus across multiple skill domains valued by the labor market but reduces the competitive edge in skills that matter the most. An alternative interpretation is that catering to the general labor market produce graduates who are well matched to entry-level positions but less able to advance, resulting in lower earnings over the 1-9 year horizon we observe. These two mechanisms are not mutually exclusive and both predict the same sign of the effect.

7.5 Which Skills Matter? Earnings Effects by Skill Cluster

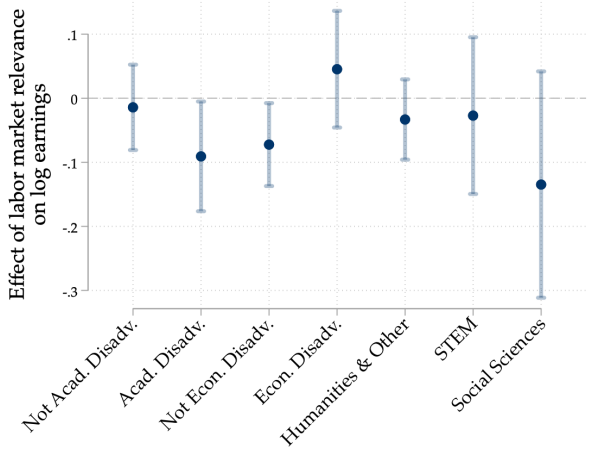
The aggregate LMR index Φ_s compresses the content of all 29 Lightcast skill clusters into a single number, weighted by local labor-market demand. The aggregate results reveal that higher LMR

Figure 4: Effects of Course Content on Log Earnings, by Subgroup

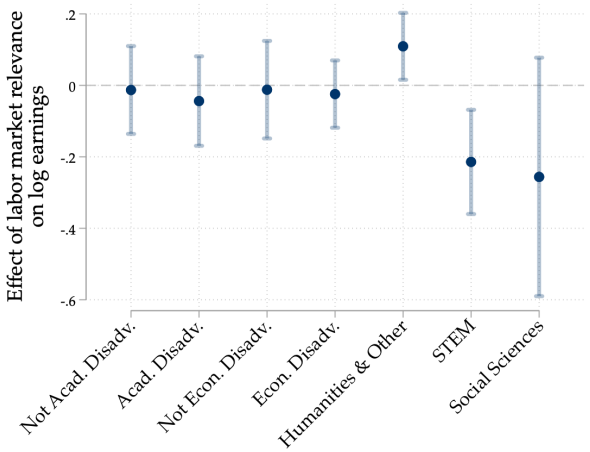
(a) Academic relevance, academic students



(b) Labor-market relevance, academic students



(c) Labor-market relevance, vocational students

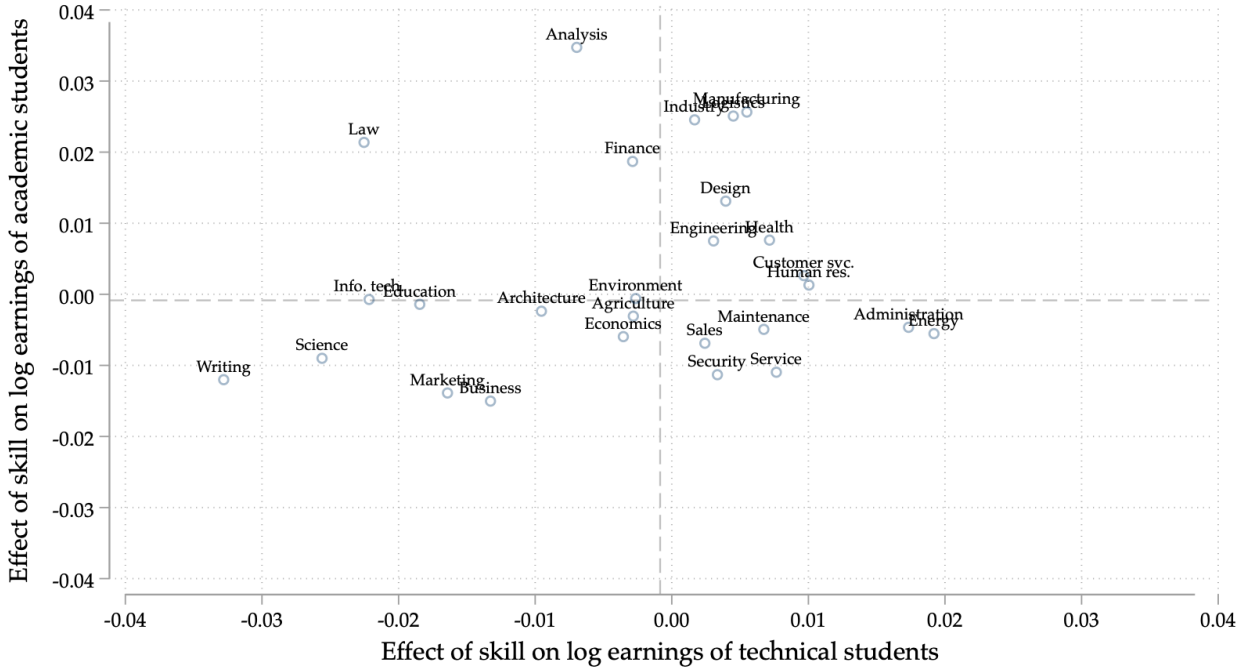


Notes: Each panel plots OLS estimates and 95% confidence intervals of the parameter β in equation (2), estimated separately using the indicated outcome as the dependent variable and on the subsamples indicated on the horizontal axis. Subsamples are defined by academic disadvantage status (not academically disadvantaged, academically disadvantaged), economic disadvantage status (not economically disadvantaged, economically disadvantaged), and broad field of study (Humanities & Other, STEM, Social Sciences). Panel (a) shows effects of standardized average academic relevance for students in academic programs. Panels (b) and (c) show effects of standardized average labor-market relevance for students in academic and vocational programs, respectively. All specifications control for lagged academic relevance; college-by-major-by-cohort fixed effects; indicators for instructors having a Ph.D. and for having worked at a four-year institution; and for the following student characteristics: gender, ethnicity, age, economic disadvantage, high school test score, the high school share of free- or reduced-price-lunch students, under-represented minorities, student-staff ratio, county median income, and locale. We additionally control for a quadratic of potential experience. Confidence intervals are constructed using standard errors clustered at the student level.

harms academic students and leaves vocational students' earnings unchanged on average. However, these patterns may conceal meaningful differences across skill types. We now decompose the

analysis by estimating, for each of the 29 clusters separately, the effect of that cluster's presence in a student's courses on post-graduation log earnings. Figure 5 plots the results as a scatter, with the earnings effect for vocational students on the horizontal axis and the earnings effect for academic students on the vertical axis. Each point is a skill cluster, labeled by name.

Figure 5: Effects of Specific Skill Clusters on Log Earnings of Academic and Vocational Students



Notes: Each point represents the effect of a one-sd increase in the average density of one of the 29 Lightcast skill clusters. The horizontal axis plots the OLS estimate of the effect of that cluster's demand-weighted density in a student's courses on log quarterly earnings, estimated on students in vocational programs. The vertical axis plots the analogous estimate for students in academic programs. Both regressions follow the lagged-content specification in equation (2) and include the same controls as Tables 4 and 6, respectively. One observation is a student-quarter; standard errors are clustered at the student level.

The scatter reveals a rich and interpretable structure. Four regions are of particular interest.

Skills that benefit both program types (upper-right quadrant). A cluster of applied technical and production-oriented skills — Manufacturing, Industrial, Design, Engineering, Health, Customer Service, and Human Resources — yield positive or near-positive earnings effects for both academic and vocational students. These are broadly valued competencies in growing sectors of the Texas economy. Their positive effect for academic students is notable: exposure to production-oriented course content does not harm academic trajectories and appears to complement the academic preparation measured by our other content dimension.

Skills that benefit academic but not vocational students (upper-left quadrant). Analysis, Law, and Finance have the largest positive effects for academic students among clusters with negative effects for vocational students (approximately 3.4, 2.1 and 2.0 log points, respectively, on the vertical axis). This pattern is consistent with evidence on the growing returns to cognitive skills (Deming, 2017). It also reflects the career paths of academic-program completers: those who transfer, obtain bachelor's degrees, and enter professional labor markets in law, finance, and related fields reap large returns to having encountered this content at the community college level. Vocational program completers, in contrast, are typically hired into operational or trade positions where analytical, legal, and financial knowledge confers little wage advantage.

Skills that benefit vocational but not academic students (lower-right quadrant). In this set, administration stands out, with an earnings effect of approximately 1.7 log points for vocational students while being essentially neutral for academic students. Administrative competencies such as scheduling, record keeping, and office operations are directly rewarded in the entry-level positions that vocational graduates typically fill, but do not contribute to the academic-track career premium that accrues primarily through credentials and subject-matter depth. Energy also substantially increases earnings for vocational students, but this might be specific to the Texas-context.

Skills harmful to both program types (lower-left quadrant). The most striking finding in Figure 5 is the cluster of skill types with negative earnings effects for both academic and vocational students: Writing, Science, Marketing, and Business all sit in the lower-left quadrant. In the Lightcast taxonomy, these labels refer to specific occupational skill clusters — technical writing, life science and medical laboratory skills, marketing operations, and business administration workflows — rather than the broad competencies their names might suggest. Heavy course-level emphasis on these narrow occupational clusters may crowd out preparation that would be more valuable on either pathway: for academic students, the crowded-out content is likely the academic rigor that supports transfer and degree completion; for vocational students, it is the applied production and operational skills that generate wage premiums at hire. These clusters thus represent a form of misallocation that is doubly costly, harming students on both tracks.

8 Conclusion

This paper studies how course content shapes the outcomes of community college students. Using 605,913 syllabi from 40 Texas community colleges linked to administrative records on degree attainment, transfer, and earnings, we construct course-level measures of academic relevance (AR)

and labor-market relevance (LMR) and exploit two sources of plausibly exogenous variation — last-minute syllabus updates unobservable to students at enrollment, and geographic variation in the content of nearby colleges — to identify their effects. Both strategies yield a consistent picture: what is taught in community college courses is an important determinant of students' educational and economic trajectories, with effects that depend sharply on the type of program students are enrolled in.

These findings speak to a broader point about community college quality. Unlike class size, instructor ratings, or graduation rates — all of which receive institutional and policy attention — the academic orientation of course syllabi is difficult for students to observe before enrollment and are rarely featured in accountability metrics. Our results show that it nonetheless shapes whether students complete degrees and what they earn years later. This invisibility has a practical consequence: institutions and instructors face ongoing pressure to demonstrate vocational relevance, and for the majority of students enrolled in academic programs, yielding to this pressure comes at a real cost to the outcomes those students most value.

Several questions remain open. Our earnings window covers one to nine years post-graduation; longer horizons may reveal additional returns to academic rigor, particularly for students whose credential premium accrues over a full career. The mechanisms behind the large negative LMR effects for STEM vocational students are not fully resolved: does broad labor market relevance crowd out specific skill clusters that matter the most for vocational students, or does it lead students to sort into lower advancement entry-level occupations? And while Texas is a large and diverse state, whether these patterns generalize to other labor markets and institutional contexts is an important open question. Perhaps most consequentially, our results raise a policy question we cannot yet answer: can targeted interventions shift course content in productivity-enhancing directions at scale? Answering this will require both more granular evidence on how content decisions are made and quasi-experimental variations in relevant institutional and state policies.

References

Acemoglu, D. and D. H. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 1043–1171. Elsevier.

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics 4*, 185–223.

American Association of Community Colleges (2023). Fast facts 2023.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Bailey, T., S. S. Jaggars, and D. Jenkins (2015). *Redesigning America's Community Colleges: A Clearer Path to Student Success*. Harvard University Press.

Bettinger, E. P. and B. T. Long (2010). Does cheaper mean better? the impact of using adjunct instructors on student outcomes. *Review of Economics and Statistics 92*(3), 598–613.

Biasi, B. and S. Ma (2022). The education-innovation gap. Technical report, National Bureau of Economic Research.

Biasi, B. and S. Ma (2026). Frontier knowledge in college and student success.

Brunello, G. and D. Checchi (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy 22*(52), 782–861.

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

Chau, H., S. H. Bana, B. Bouvier, and M. R. Frank (2023). Connecting higher education to workplace activities and earnings. *Plos one 18*(3), e0282323.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020). Income segregation and intergenerational mobility across colleges in the United States. *Quarterly Journal of Economics 135*(3), 1567–1633.

Cohen, A. M., F. B. Brawer, and C. B. Kisker (2014). *The American Community College* (6th ed.). Jossey-Bass.

College Board (2023). Trends in college pricing and student aid 2023.

Deming, D. and S. Dynarski (2009). Into college, out of poverty? policies to increase the postsecondary attainment of the poor. *NBER Working Paper No. 15387*.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics 132*(4), 1593–1640.

Deming, D. J. and L. B. Kahn (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics 36*(S1), 337–369.

Doyle, W. R. (2009). The effect of community college enrollment on bachelor's degree completion. *Economics of Education Review 28*(2), 199–206.

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

Grosz, M. (2020). The returns to a large community college program: Evidence from admissions lotteries. *American Economic Journal: Economic Policy 12*(1), 226–253.

Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017). General education, vocational education, and labor-market outcomes over the life-cycle. *Journal of Human Resources 52*(1), 48–87.

Hershbein, B. and L. B. Kahn (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review 108*(7), 1737–1772.

Jacobson, L., R. J. LaLonde, and D. G. Sullivan (2005). Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics 125*(1-2), 271–304.

Jepsen, C., K. Troske, and P. Coomes (2014). The labor-market returns to community college degrees, diplomas, and certificates. *Journal of Labor Economics 32*(1), 95–121.

Kane, T. J. and C. E. Rouse (1995). Labor market returns to two- and four-year college: Is a credit a credit and do degrees matter? *American Economic Review 85*(3), 600–614.

Mountjoy, J. (2022). Community colleges and upward mobility. *American Economic Review 112*(8), 2580–2630.

National Center for Education Statistics (2023). Digest of education statistics 2023.

Nowogrodzki, A. (2016). Mining the secrets of college syllabuses. *Nature 539*(7627), 125–126.

Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics 13*(2), 217–224.

Stevens, A. H., M. Kurlaender, and M. Grosz (2019). Career technical education and labor market outcomes: Evidence from California community colleges. *Journal of Human Resources 54*(4), 986–1036.

Texas Higher Education Coordinating Board (2023a). Texas core curriculum. Policy Documents and Guidelines.

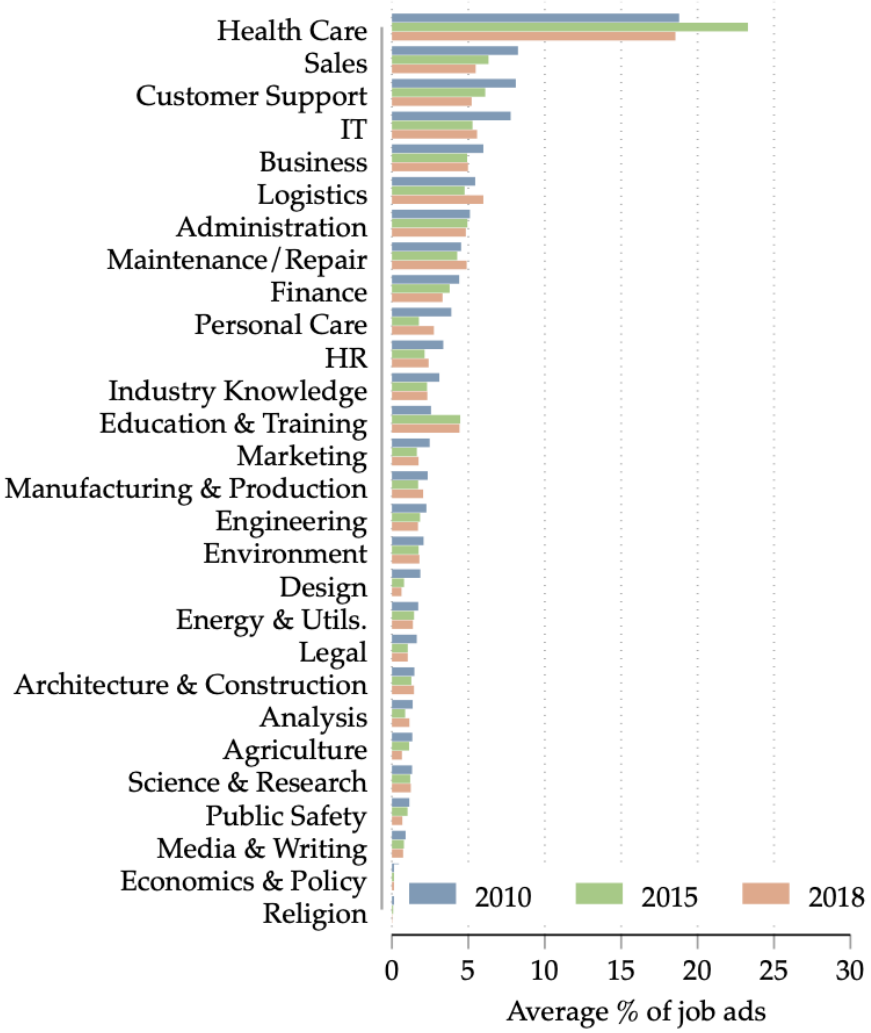
Texas Higher Education Coordinating Board (2023b). Texas public higher education almanac 2023.

Appendix

For online publication only

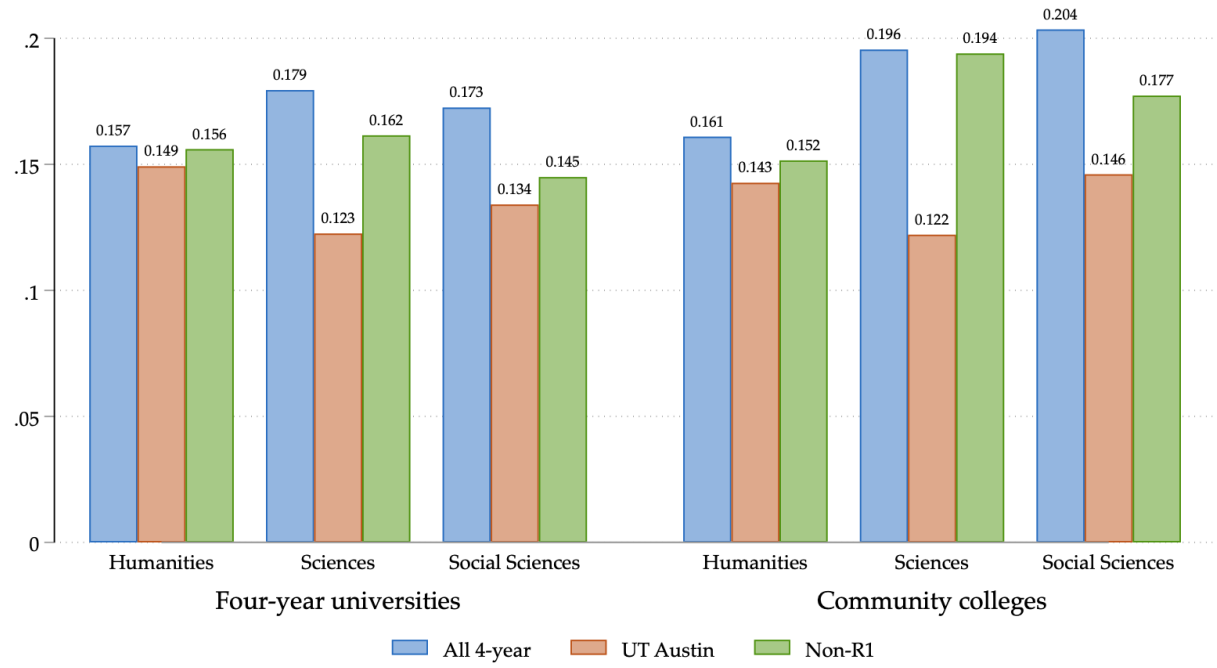
Appendix A Additional Tables and Figures

Figure A1: Skill Demand Across Job Ads



Notes: Average share of jobs ads requiring skills in each family, across counties, in 2010, 2015, and 2019.

Figure A2: Academic Relevance, by Institution Type and Field

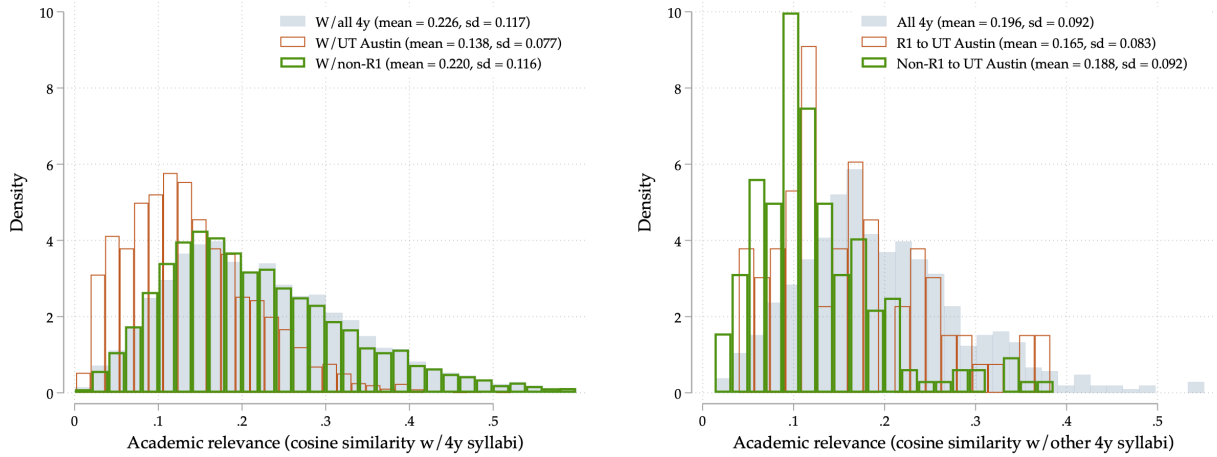


Notes: Histogram of academic relevance across courses at community colleges and at four-year institutions in our sample, by macro-field. Academic relevance is defined as a course syllabus's average cosine similarity with 4-year courses with the same code within the Texas Core Curriculum.

Figure A3: Distribution of Maximum Academic Relevance, by Institution Type

(a) Community colleges

(b) Four-year institutions



Notes: Histogram of maximum academic relevance across courses at community colleges and at four-year institutions in our sample, overall and by type of comparison institution. Academic relevance is defined as a course syllabus's maximum cosine similarity with 4-year courses with the same code within the Texas Core Curriculum.

Appendix B Syllabi Data Details

The vectorization procedure follows Biasi and Ma (2026). To clean the text of each syllabus, we proceed as follows:

(i) We first convert all PDF files of syllabi to texts using PyMuPDF.[12]

(ii) We use the Unidecode Python Package[13] to convert Unicode text into ASCII text. This includes legacy code that does not support Unicode, non-Roman names on a US keyboard, and ASCII approximations for symbols and non-Latin alphabets.

(iii) We remove browser information, often present in the header of a syllabus, by searching for keywords such as "Internet Explorer", "Newer Browser", "JavaScript Enabled", "Cookies Are", "Download Info", "Login", "Log In", "Print", and "Search".

To transform each document into a word vector, we use as a dictionary the set of all unique unigrams, bigrams, and trigrams ever appearing as keywords in OpenAlex, an openly licensed (CC0) catalog of global research. Maintained by OurResearch, OpenAlex indexes over 240 million scholarly works with approximately 50,000 new works added daily, drawing from multiple sources such as Crossref, PubMed, institutional repositories (e.g., arXiv), and the legacy Microsoft Academic Graph. database of approximately 107 million academic articles.

From each syllabus, we extract N-grams (N $\in$ [1, 3]) from each document and retain only those matching a dictionary entry. Each document is then represented as a term-frequency vector, weighted using a backward-looking TF-IDF (TFBIDF) scheme (Biasi and Ma, 2022, 2026). Specifically, the inverse document frequency for term w at time t is computed only over documents published prior to t, so that newly emerging terms receive appropriately high weights even before they become widespread. We compute textual similarity between documents as the cosine similarity of their TFBIDF-weighted vectors.

[12] https://pymupdf.readthedocs.io/en/latest/
[13] https://pypi.org/project/Unidecode/

Appendix C Institutional Evidence on Syllabus Visibility at Community Colleges

This appendix documents the institutional basis for a key premise of our identification strategy: that students cannot observe the *updated, current-term syllabus* when they select courses. The identifying assumption does not require that students can never locate *any* syllabus—prior-term syllabi are often accessible—but rather that last-minute content updates are not systematically visible at the time of initial enrollment. The institutional evidence assembled here supports this claim across the colleges in our sample.

C.1 Statewide Framework

The starting point is Texas's HB 2504 regime, codified in Texas Education Code Sec. 51.974.[14] The statute requires Texas public institutions to make undergraduate course syllabi and instructor curricula vitae publicly available online. The relevant statewide implementing rule requires public posting *within seven days of the first class day*, not by the opening of registration.[15] Because registration typically opens well before the semester starts, the statewide rule itself implies that students are not guaranteed access to the updated syllabus when they select courses.

The statewide rule does not preclude individual instructors from posting early. It does establish the narrower and empirically relevant point that updated-syllabus visibility at registration is not required by the governing institutional regime. For several colleges in our sample, college-specific materials restate this timing directly. For others, official syllabus repositories exist without equally explicit timing language. Table C1 summarizes the evidence by tier.

C.2 Tier 1: Colleges with Direct Institution-Specific Timing Evidence

Angelina College. Angelina College's official HB 2504 page states that the college provides public access to syllabi and instructor curricula vitae "within one week after the first day of class."[16] Students registering before the semester begins therefore cannot observe the updated current-term syllabus.

[14] https://statutes.capitol.texas.gov/Docs/ED/htm/ED.51.htm#51.974.

[15] https://texreg.sos.state.tx.us/public/readtac$ext.ViewTAC?tac_view=5&ti=19&pt=1&ch=4&sch=A&rl=Y.

[16] https://www.angelina.edu/facultyandstaff/instructional-resources-faculty/house-bill-2504/.

Table C1: Summary of Institutional Evidence on Syllabus Visibility

Evidence tier	Interpretation	Number of colleges	Representative examples
Tier 1: explicit institution-specific timing language	Students could not rely on observing the updated syllabus at registration	7	Austin CC, Angelina, Cisco, Coastal Bend, Hill, Panola, San Jacinto
Tier 2: institution-specific syllabus page; timing not explicitly stated	Updated-syllabus visibility at registration not guaranteed by institutional policy	12	Brazosport, Collin, El Paso CC, Houston CC, Navarro, Northeast Texas, South Texas, Tyler JC, Vernon
Tier 3: statewide baseline	Statewide HB 2504 regime supports "not guaranteed" claim	20	Alvin, Amarillo, Blinn, Central Texas, Del Mar, Galveston, McLennan, Odessa, Weatherford

Notes: Colleges are classified by the strength of available institutional documentation. Tier 3 does not mean colleges lack public syllabi; it means that college-specific timing language beyond the statewide baseline was not identified.

Austin Community College. Austin Community College's faculty handbook states that a syllabus and current CV must be posted "no later than the seventh day after the first day of class."[17] Mandated public availability is thus pushed past the point at which most students select courses.

Cisco College. Cisco College's HB 2504 page states that public course information is posted "by no later than the seventh day after the first day of class."[18]

Coastal Bend College. Coastal Bend College's HB 2504 page states that syllabus and instructor information are accessible online "by no later than the seventh day after the first day of class."[19]

Hill College. Hill College's faculty handbook states that the complete syllabus is due "by the 7th class day."[20]

Panola College. Panola College's HB 2504 compliance page states that public access is provided "by the seventh day after the first day of class" and notes that this structure has been in place

[17] https://sites.austincc.edu/faculty-handbook/teaching/requirements-and-guidelines/course-syllabus/.

[18] https://www.cisco.edu/house-bill-2504-public-access-to-course-information/.

[19] https://www.coastalbend.edu/House-Bill-2504/.

[20] https://www.hillcollege.edu/About_HC/administration/faculty_handbook/course-syllabus-guidelines.html.

since January 2011.[21] Panola thus provides both strong current evidence and continuity into the beginning of the 2011–2022 sample window.

San Jacinto College. San Jacinto College's faculty handbook states that syllabi and CVs must be available "by the seventh day after the first day of class."[22]

C.3 Tier 2: Colleges with Supportive Indirect Evidence

Collin College. Collin College maintains a public-access page for course information consistent with the HB 2504 framework.[23] Institutional materials do not specify that updated syllabi must be posted only after classes begin, so the appropriate claim is the narrower one that students were not shown to have a guaranteed right to inspect the updated syllabus when registration opened.

Dallas College. Dallas College (the product of the reorganization of the Dallas County Community College District) maintains a public academic calendar that clearly shows term structure and later-start options.[24] Institutional materials do not include a syllabus-timing statement of the strength found in Tier 1 colleges; the evidence accordingly supports only the narrower claim that updated-syllabus visibility at initial course selection is not established by institutional policy.

El Paso Community College. EPCC maintains formal syllabus procedures documented in administrative policy.[25] A formal institutional procedure from 2011 establishes that structured syllabus disclosure existed at the beginning of the sample period. Institutional materials do not, however, specify that public visibility begins only at the seventh class day; EPCC is therefore best treated as supportive of the narrower "not guaranteed" claim.

Houston Community College. Houston Community College maintains a public syllabus search portal.[26] The existence of a searchable repository does not establish that the updated current-term syllabus must be available when students first register; the evidence is accordingly consistent with but does not directly establish the identifying assumption.

[21] https://www.panola.edu/panola-college-public-access-to-course-information/.
[22] https://www.sanjac.edu/about/faculty-and-staff/resources/faculty-handbook/college-operations/index.html.
[23] https://www.collin.edu/academics/programs/publicaccess.html.
[24] https://www.dallascollege.edu/resources/academic-calendar/pages/default.aspx.
[25] https://www.epcc.edu/AboutEPCC/board-of-trustees/board-policy/administrative-procedures/procedures-general-institution; https://www.epcc.edu/AboutEPCC/board-of-trustees/board-policy/administrative-procedures/Documents/GM/6_31_01.pdf.
[26] https://learning.hccs.edu/faculty/syllabus-search.

Brazosport College, Navarro College, Northeast Texas Community College, South Texas College, Temple College, Tyler Junior College, and Vernon College. Each of these colleges maintains an institution-specific public-access, HB 2504, handbook, or syllabus page consistent with public syllabus availability, without providing the explicit timing language present at Tier 1 institutions.[27] For these institutions, the available evidence supports the claim that students were not guaranteed access to the updated syllabus when registration opened, consistent with the statewide HB 2504 baseline.

C.4 Tier 3: Statewide Baseline

For the remaining colleges in the sample—including Alvin Community College, Amarillo College, Blinn College, Central Texas College, College of the Mainland, Del Mar College, Frank Phillips College, Galveston College, Howard College, Kilgore College, Laredo College, McLennan Community College, Midland College, Odessa College, Ranger College, Southwest Texas Junior College, Texas Southmost College, Trinity Valley Community College, Weatherford College, Western Texas College, and Wharton County Junior College—the available institutional documentation does not go beyond the statewide HB 2504 framework. As Texas public community colleges, these institutions are subject to the same statewide regime that places mandated syllabus posting after the start of the term, which supports the claim that students were not guaranteed access to the updated syllabus at the outset of registration.

C.5 Enrollment Adjustment After Classes Begin

A complementary institutional fact relevant for interpretation is that students at most Texas community colleges can revise course choices after classes start. Austin Community College's student calendar states that students may add classes through the first business day of the term and may drop with later refund deadlines.[28] Dallas College's calendar distinguishes regular and flex terms, leaving open enrollment opportunities in later-start courses after the standard semester begins.[29] Northeast Texas Community College's and Vernon College's calendars similarly include multiple

[27] https://www.brazosport.edu/faculty-staff/instructional-technology/distance-education/house-bill-2504-faculty-info; https://www.navarrocollege.edu/academics/course-syllabi.html; https://www.ntcc.edu/academic-affairs/house-bill-2504; https://www.southtexascollege.edu/academics/pdf/employee-handbook.pdf; https://www.templejc.edu/about-tc/departments/hb2504-course-information; https://www.tjc.edu/info/200344/public_access_to_course_information; https://syllabi.vernoncollege.edu/.

[28] https://students.austincc.edu/registration/academic-calendar/.

[29] https://www.dallascollege.edu/resources/academic-calendar/pages/default.aspx.

start dates and later-start sessions.[30] This structure is consistent with our empirical design: while students cannot observe last-minute syllabus updates before initial enrollment, they retain the ability to adjust their schedule once the updated syllabus becomes available.

Taken together, the evidence establishes that across Texas public community colleges, students were not guaranteed access to the updated current-term syllabus when selecting courses. For seven colleges in our sample, institutional materials state this timing directly. For the remainder, the statewide HB 2504 framework and institution-specific public-access policies are consistent with the same conclusion. The modal pattern across the sample is for public access to the updated syllabus to begin after course selection is already underway, and often only after classes begin.

[30] https://www.ntcc.edu/academic-calendar; https://www.vernoncollege.edu/payment-and-registration-dates.