Who Gets to Use Their College Skills? Skill-Job Alignment and Earnings Inequality Among College Graduates

Aleksei Opacic*

Harvard University

Abstract

A growing literature links inequalities among college graduates to education–job alignment. But because existing measures infer alignment from common credential-to-occupation pathways, rather than directly measuring students' skills, they miss broader forms of skill transfer across occupations. How prevalent are these broader forms of skill utilization, and how much do they matter for earnings inequalities? I address these questions by developing a framework of transferable human capital and measuring skill-task match: the alignment between graduates' transcript-derived skill profiles and their job tasks. Empirically, I construct the first individual-level linkage between college course-catalog records and students' jobs using administrative transcript and employment data from a U.S. state. Leveraging this dataset, I document three findings. First, there is substantial variation in the extent to which graduates use their skills at work. Second, graduates who do work in skill-aligned jobs earn significantly more than their peers, even within the same occupation. Third, there are large demographic disparities in access to skill-aligned work; critically, these disparities account for a full quarter of the Black–White earnings gap. Common education-to-work pathways are one route into skill-utilizing work, but much of higher education's economic value—and inequalities in access to that value—depends on skill transfer across occupational boundaries.

---

*Department of Sociology, Harvard University, aopacic@g.harvard.edu. I am grateful to Clem Aeppli, Siwei Cheng, Samuel Donahue, Christina Ciocca Eller, Keitaro Okura, Nathan Wilmers, Chris Winship, and Xiang Zhou for helpful comments on previous drafts. The analyses and conclusions presented in this work are solely those of the author and do not reflect the views or opinions of the state's Employment or Education Departments. This research was partly supported by a Russell Sage Foundation Dissertation Grant (RSF Grant#: 2401-47047), grants from the Harvard Institute for Quantitative Social Science, the Harvard Center for Population Studies, a Harvard Kennedy School Stone Research Grant, as well as a Linda G. Hammett Ory Fellowship.

1 Introduction

To what extent do college graduates use their skills in the labor market and with what consequences for earnings inequalities? A college degree remains one of the central routes to economic opportunity in the United States, but degree holders do not share equally in its rewards. To explain earnings inequality among college graduates, a large literature across sociology and economics has revealed the importance of qualitative dimensions of a person's education – from college quality (Bowen and Bok, 1998; Brand and Halaby, 2006; Witteveen and Attewell, 2017a), to field of study (Gerber and Cheung, 2008; Altonji et al., 2014), and, more recently, to the specific courses students take within college (Han et al., 2023; Kizilcec et al., 2023). But the labor-market value of education also depends on whether graduates enter jobs that allow them to use their skills – an insight that has motivated an interdisciplinary literature on education–job match: the alignment between a worker's schooling and their job (Kalleberg, 2008; Witte and Kalleberg, 1995; de Werfhorst, 2002; Wolbers, 2003; Robst, 2007). Match, so-defined, is an important driver of earnings, especially when graduates follow well-defined pathways between their field of study and the labor market (DiPrete et al., 2017; Bol et al., 2019).

Yet existing measures of education–job match capture just one form of skill–job alignment. In much of this work, graduates are treated as matched when they enter common occupational destinations for their field of study, and mismatched when they work outside those destinations. Such a credential-to-occupation conception of match is useful for fields that prepare students for a small set of licensed or occupation-specific jobs, such as nursing, accounting, or engineering, but is less suited to the U.S. higher-education system, where liberal-arts curricular traditions and comparatively weak coordination between employers and educational institutions mean that many degrees do not operate as direct pipelines into specific occupations. At the same time, contemporary labor markets increasingly favor not just technical and occupation-specific expertise but higher-order skills such as problem solving, quantitative analysis, written communication, and analytical reasoning (AACU, 2010; Autor et al., 2003; Moss-Pech, 2025). If employers increasingly demand general-education skills in addition to occupationally specific training, then a central question is whether higher education's broad and transferable skills are actually mobilized in the labor market and with what consequence. How prevalent are these broader forms of skill-job alignment, and how much do they matter for earnings among college graduates?

To answer these questions, I develop a framework of transferable human capital. The framework begins from the premise that occupations are not simply discrete categories, but bundles of tasks, skills, and responsibilities (Liu and Grusky, 2013; Han and Cheng, 2026; York et al., 2025). Because task content overlaps across occupations, workers can transfer skills across occupations that rely on similar tasks (Gathmann and Schönberg, 2010; Yamaguchi, 2012). I extend this insight to higher education by conceptualizing college coursework as developing task-relevant capacities that may transfer across jobs drawing on similar underlying abilities. I use the term *skill-task match* to describe the alignment between the skills graduates develop through coursework and the tasks they are required to perform at work. Compared with credential-to-occupation measures, which infer alignment from common pathways between credentials and occupations (for example, a nursing major who becomes a nurse or a physics major who becomes a physicist), I directly measure the skills education imparts, and whether those skills align with the tasks graduates perform at work. For example, a physics major who becomes a data scientist or a sociology graduate who becomes a market researcher may not be following a common pathway, but nevertheless uses their skills at work. My approach recovers these broader forms of education-job alignment.

Empirically, I construct skill-task match by locating courses and occupations in a shared task space. I first use workplace activity data to create a low-dimensional representation of occupational tasks, where proximity captures tasks that draw on similar skills. I then use course-catalog text to infer the task-relevant content of each course and aggregate these course-level measures into student transcript profiles. Skill-task match is defined as the similarity between a graduate's transcript-derived skill profile and the task content of their post-college occupation.[1] I study skill-task match among BA degree holders in a U.S. state, using a new panel dataset that links administrative postsecondary transcripts to employment records, and merge this with textual course-catalog data. The resulting dataset is unique in its ability to uncover the education-to-work transition at the level at which skill acquisition and deployment actually occurs. Rather than observing only declared majors, I am able to link students' course-taking behavior to course-catalog text that can be systematically compared with jobs' task demands. This structure allows me to, for the first time, ask whether the substantive content of students' coursework aligns with the task demands of the occupations they later enter.[2]

The analyses yield three main findings. First, skill-task match varies substantially within majors, occupations, and common credential-to-occupation pathways, indicat-

[1] I observe individuals' occupations in the administrative data, rather than their jobs. This limitation is important, and motivates future data linkages that identify task usage at a more micro level. In general, however, defining job tasks at the level of the occupation rather than job likely leads me to under-estimate true skill-task match: some individuals not matched to the centroid vector of a given occupation may nevertheless be matched to a specific job position within the occupational category. I return to this limitation in the discussion section.

[2] The framework is not specific to BA degrees. In principle, skill-task match could be measured for any postsecondary credential. In this paper, I focus empirically on BA graduates for reasons of measurement feasibility, as discussed in the Data section.

ing that common match categories obscure important variation in education–job alignment. Second, skill-task match predicts earnings net of measured skills, credentials and occupation. Much of what appears to be the earnings return to common credential-to-occupation matching is attributable to a more general and continuous measure of skill-task alignment. Third, I find significant race- and income-based disparities in task matching; skill-job misalignment is an important driver of earnings inequality between these groups. Importantly, these disparities are invisible under credential-based definitions of match, implying that they emerge outside, rather than along, common school-to-work pathways.

This paper shows that access to skill-aligned work beyond occupation-specific pathways is a motor of earnings inequality among college graduates. In so doing, it traces the roots of unequal returns to college past the well-documented educational inputs that shape earnings – such as college quality, field of study, and course-taking – to the relational position between graduates' skills and their work. As technological transformations continue apace in altering the task structure of work, and as labor market demand for new combinations of analytic and technical capacities replaces old skill demands, the ability to transfer skills into new jobs and occupations may increasingly determine who benefits – and who loses out – from these transformations.

2 Education–Job Match and the Limits of Credential Pathways

Sociologists have long treated occupational allocation as a central inequality-producing mechanism (Blau and Duncan, 1967; Breen and Jonsson, 2005; Breen et al., 2016). But the jobs graduates obtain matter not only because occupations hold different positions in the stratification order. A long tradition in sociology and economics treats the labor-market value of education as relational rather than intrinsic: the value of a person's schooling depends partly on its fit with the tasks and requirements of that person's job. While human capital theory begins from the premise that schooling raises earnings by developing productive capacities (Becker, 1962; Mincer, 1974), job-matching and assignment models have argued that those capacities generate returns only insofar as they can be deployed in work. In these approaches, earnings depend not only on worker attributes or job attributes in isolation, but on the fit between workers and positions (Jovanovic, 1979; Sattinger, 1993). Occupational sorting is therefore a process through which educational investments are either activated or underused.

Sociological research on educational mismatch has developed this relational insight in two ways. Early work examined vertical mismatch (whether workers have more or less education than their jobs require), documenting penalties to over-education and under-utilized schooling (Kalleberg, 2008; Leuven and Oosterbeek, 2011). But while vertical mismatch captures an important form of failed translation between schooling and work, it says less about whether the specific content of a worker's education is useful for the tasks they perform. For this reason, researchers have expanded the study of match to a horizontal perspective: the effects of pursuing an occupation that aligns with a worker's credential (de Werfhorst, 2002; Wolbers, 2003; Bol et al., 2019; Li and Lu, 2023; Lu et al., 2025). Early studies, using researcher-defined classifications of aligned occupations, found that graduates often earn more when they work in jobs connected to their educational field (de Werfhorst, 2002; Wolbers, 2003; Bol et al., 2019), although the size of the match premium has been disputed (Béduwé and Giret, 2011).

More recent work has extended this approach by using more objective classifications of alignment - defining match as following an occupational pathway that is common or expected given a worker's credential – and focusing on the moderating role of school-to-work linkages: the degree to which particular credentials channel graduates into a small set of occupational destinations (Roksa and Levey, 2010; DiPrete et al., 2017; Forster and Bol, 2018; Bol et al., 2019). This literature shows that credential-to-occupation matching can be especially valuable when programs are linked to a small set of occupations. Graduates from these programs (such as medicine or nursing programs) who follow their expected occupational pathway are able to more immediately leverage their skills, and further benefit from credentialing, licensing, and other professional closure devices that regulate access to the occupation.

These prior approaches capture important forms of match, especially where credentials provide occupation-specific skills, but there are three reasons to broaden their scope. First, many four-year degrees in the United States do not have an obvious occupational match. Fields such as the humanities, social sciences, and many natural sciences often prepare graduates for a range of possible destinations rather than a single expected job. American vocational qualifications are also not nationally standardized to the extent they are in other industrialized countries, while liberal arts curriculums mean that students may complete different course sequences and acquire different skill combinations even within the same major (Han et al., 2023; Kizilcec et al., 2023). For this sizeable number of graduates, the concept of an expected pathway may break down: working outside an expected occupation does not necessarily mean that a graduate is failing to utilize their college-acquired skills.

Second, credential-to-occupation measures risk equating educational relevance with following occupationally specific pathways. Yet many capacities demanded in the contemporary labor market are not the purview of a single major or professional program. Employers increasingly value nonroutine analytic skills such as quantitative, analytical, interactive, and communication-intensive capacities alongside more explicitly vocational preparation (Autor et al., 2003; Deming and Kahn, 2018; AACU, 2010; Moss-Pech, 2025). These capacities appear in different combinations across occupations and, importantly, can be developed through multiple curricular routes. A physics graduate may not be employed as a physicist, for example, but could nonetheless employ their quantitative and computational skills across a variety of jobs. Relatedly, a sociology graduate may not be trained for one specific occupation in the way that a nursing or engineering graduate is, but coursework in statistics, organizations, and social research may prepare that graduate for work in policy analysis, market research, or data-oriented organizational roles. Recent qualitative work similarly suggests that liberal arts graduates often apply specific degree-acquired skills, such as professional writing and research, to substantive work tasks — and sometimes to a greater extent than their peers in 'practical' majors (Moss-Pech, 2025). If higher education prepares students not only for specific occupations through vocational training but also for broader forms of analytic, quantitative, interactive, and organizational work, then educational relevance cannot be inferred only from whether graduates enter common credential-to-occupation pathways.

Third, existing measures infer match from observed pathways instead of measuring educational content directly. Graduates are treated as matched when they enter occupations commonly associated with their field, but the skills imparted by their coursework and the task demands of their jobs are not directly observed. This inference is plausible in highly occupation-specific fields such as medicine, nursing or engineering, where the credential, curriculum, and job are tightly linked. But pathway-based measures can miss transferable skill use outside common destinations and can also overstate alignment when graduates enter common destinations that do not actually draw on their specific coursework.

Together, these limitations leave open a broader question: how prevalent and economically important are forms of education–job alignment that occur outside common credential-to-occupation pathways? Answering this question requires moving beyond nominal categories, such as degree level, field of study, and modal occupational destination, toward a direct measure of the alignment between educational content and job tasks.

3 Skill-task match

3.1 Overview of approach

To answer this question, I conceptualize college-acquired skills as portable investments that can be activated across a range of labor-market settings. College courses do not simply train students for specific occupations or discrete workplace tasks, but instead develop task-relevant capacities. These capacities are transferable across jobs that draw on similar underlying skills. From this perspective, a college course may provide preparation for two jobs not only because those jobs involve identical tasks, but because their tasks require related forms of knowledge and skill. For example, a physics course may prepare students for both a scientific role that requires detection of physical phenomena and reporting experimental results and a data science role that involves analyzing data and performing simulations. These jobs involve different substantive domains and different workplace tasks, but both draw on overlapping capacities in statistical reasoning, data analysis, and communicating quantitative evidence. In this sense, coursework in physics may provide preparation for both. I therefore characterize college coursework in terms of the latent task domains with which its content is associated. Alignment between education and jobs can then be measured along a continuous spectrum: as the degree to which the capacities developed through a graduate's coursework correspond to the capacities required by their job. I refer to this form of alignment as *skill-task match*.

This approach builds directly on relational perspectives on the labor market, which understand jobs not as nominal categories, but as positions defined by overlapping bundles of tasks, skills, and responsibilities (Cheng et al., 2019; Deming, 2017; Liu and Grusky, 2013; Han and Cheng, 2026). I extend this perspective to education by locating courses and transcript trajectories in the same task space used to characterize jobs. It also builds on earlier subjective measures of match, which captured a broader definition of labor-market alignment by asking workers whether their jobs used their education or training (Robst, 2007; Levels et al., 2014). These measures are useful because they allow match to occur anywhere in the occupational structure, rather than only along common credential-to-occupation pathways, but rely on workers' subjective assessments of fit. I improve on these earlier measures by providing a systematic measure of education–job alignment based on the proximity between transcript-derived skill profiles and job task requirements.

3.2 Portable skill investments

For decades, sociological work has conceptualized occupations as groupings of similar tasks, skills, and responsibilities rather than as entirely discrete categories (Blau and Duncan, 1967; Hauser and Warren, 1997; Hout, 1984). Data sources, such as the Occupational Information Network (O*NET), verbatim text responses from respondents (Martín-Caughey, 2021), and, more recently, task taxonomies generated from job postings data (Han and Cheng, 2026) make it possible to characterize occupations in terms of their work content rather than only as nominal categories (Autor et al., 2008; Cheng et al., 2019; Deming, 2017; Goldin and Katz, 2008; Liu and Grusky, 2013). Characterizing jobs in terms of work content helps describe structures of inequality that broad occupational categories would obscure—for example, by explaining the growth in between-occupation wage inequality in recent decades (Liu and Grusky, 2013), and by characterizing intergenerational mobility more parsimoniously than broad labor-market categories (York et al., 2025).

Leveraging the work content of occupations enables researchers to take a *relational* perspective on the labor market. Occupations can be understood as more or less similar to one another depending on the tasks they require and the skills those tasks draw upon. Recent work formalizes this intuition by representing jobs in terms of co-occurring skill and task requirements, making it possible to measure similarity among labor-market positions rather than treating occupations as unrelated categories (Han and Cheng, 2026). This relational view also underlies theories of task-specific human capital: workers can transfer skills to new occupations when the new occupation relies on similar tasks to the old one (Gathmann and Schönberg, 2010; Yamaguchi, 2012). Figure 1 shows this intuitively using a simplified two-dimensional representation of three target occupations and a physics graduate. In this representation, the angle between vectors captures similarity in task content. A data scientist, for example, may be expected to manipulate data with statistical software and identify business problems, while a physics-related scientific worker may be required to analyze data to detect physical phenomena and report experimental results.

Data-science and physics-related occupations therefore occupy nearby regions of the task space, whereas journalism sits farther away because it relies more heavily on interviewing, reporting, narrative writing, and public communication. Importantly, this proximity can arise even when the occupations do not share exact task requirements. Data-science and physics-related occupations may contain different concrete tasks in the O*NET data, but those tasks can nevertheless draw on related capacities: quantitative reasoning, statistical modeling, computational analysis, interpretation of empirical patterns, and communication of technical evidence. In this sense, the task space captures higher-order similarity in the underlying skill requirements of work, rather than only literal overlap in observed task labels.

If human capital is portable across occupations when occupations rely on similar tasks, then college-acquired skills should be portable across occupations that require similar skills. Yet in contrast to the detailed data infrastructures for measuring what workers do on the job, a comparable infrastructure for measuring what skills students acquire in college is far less developed. As a result, the large literature on the labor-market effects of higher education has often bypassed direct measures of skill acquisition, using field of study as the primary measure of educational content (Gerber and Cheung, 2008; Altonji et al., 2012, 2014; Lovenheim and Smith, 2023; Mountjoy, 2024).

But field-of-study classifications do not, in themselves, reveal the skill content of education or its returns. Nor do they allow us to infer how closely a particular educational experience aligns with the task demands of the occupations graduates enter.

To address this gap, I apply the relational task perspective used to characterize occupations to the study of education. Rather than treating higher education primarily as a set of credentials or fields of study, I conceptualize educational investments in terms of the labor-market tasks for which they prepare students. This approach makes it possible to locate both educational experiences and occupations in a shared latent task space. In this way, I define *skill-task match* as the degree of proximity between a student's course-derived skill portfolio and the task content of the occupation they enter. To construct this measure, I treat courses as the building blocks of students' skill portfolios, inferring each student's skills from their transcript trajectory, and comparing that portfolio to the task content of their job.

Figure 1 illustrates this logic. The black arrow represents a physics major. From a credential-to-occupation perspective, this graduate's expected occupation is a Physicist; the graduate would be considered unmatched if they become a Data Scientist or Journalist. From a skill-task perspective, however, a physics major is located close not only to physicist jobs but also to data-science jobs. This is because the skillsets that a physics major imparts share task families with the data scientist: mathematical modeling, computational tools, statistical analysis, and quantitative reasoning.

This framework implies that common credential-to-occupation pathways and skill-task match are related but analytically distinct dimensions of education–work alignment. Table 1 shows a cross-tabulation of skill-task match with following a common credential-to-occupation pathway. Following a common credential-to-occupation pathway coincides with high skill-task alignment (top-left cell) when the expected occupational destination for a credential draws closely on the skills that credential imparts. This could occur, for example, in highly linked fields, where the strength of a field's linkage could imply occupation-specific schooling that a graduate then leverages in their work (although high-linkage could also imply the presence of licensing regimes without a high degree of skill specificity) (DiPrete et al., 2017; Bol et al., 2019). Meanwhile, the lower-right cell represents the clearest case of mismatch or skill underutilization: graduates who enter less common destinations that do not draw strongly on their college-acquired skills.

The advantage of a skill-task match approach is that it separates the content of education–work alignment from pursuing a common credential-to-occupation pathway. The off-diagonal cells capture cases where skill-task alignment may differ from credential-to-occupation alignment. In the top right cell, individuals pursuing a common pathway may nevertheless fail to leverage their college skills at work (for example, in less linked fields of study). The lower left cell is the key case that a credential-to-occupation alignment approach obscures. A graduate may enter an occupation that is uncommon for their field while still using skills developed through coursework – precisely because those courses develop task-relevant capacities that transfer across occupations with related task demands. A relational task-space approach therefore expands the study of alignment beyond expected credential-to-occupation trajectories.

Table 1: Conceptual relationship between common credential-to-occupation pathways and skill-task match.

	High skill-task match	**Low skill-task match**
Common credential match	Common pathway with content alignment: especially in fields with occupation-specific training.	Common pathway without strong content alignment: in fields with lower-linkage, or in high-linkage fields without occupation skill specificity.
Off-pathway	Transferable skill match: the worker departs from the common pathway but enters a job that uses transferable course-acquired skills.	Likely mismatch or underutilization: neither the common pathway nor the measured task content aligns strongly with the worker's education.

In addition to expecting skill-task match to vary substantially within conventional match categories, I expect skill-task match to be associated with earnings, even net of educational inputs and net of a worker's occupation. This core implication follows from task-specific human capital theory: workers should be more productive when the tasks required in their jobs draw on skills they have already developed (Gathmann and Schönberg, 2010; Yamaguchi, 2012; Poletaev and Robinson, 2008). Signaling models would also suggest an earnings payoff to skill-task match, though for different reasons: workers better matched to a job are able to communicate their fit for the job more clearly than workers who are not; this fit is rewarded by employers, even if match does not improve a worker's productivity directly. College coursework may therefore generate returns outside expected pathways whenever it prepares students for task bundles that are valued elsewhere in the labor market.

At the same time, the returns to transferable skill alignment may be weaker or less immediate than the returns to common credential-to-occupation match. Occupation-specific pathways often make fit highly legible: credentials, professional categories, and common hiring pipelines tell employers what a worker has been trained to do

(Spence, 1973; Arrow, 1973; Bills, 2003). They also may indicate the presence of licensing regimes that restrict supply for the occupation and bolster wages for those who enter the occupation. By contrast, transferable skills are less standardized and may be harder to verify before hire. A sociology major's training in survey design, organizational analysis, or statistical reasoning may be useful in people analytics, but that fit is less categorically obvious than the fit between a nursing degree and nursing work. The absence of licensing and professional institutions may also lead to lower returns to alignment for those outside common credential-to-occupation pathways. As a result, broader skill-task match should predict earnings, but its effects may be smaller overall, and depend on whether workers and employers can identify, communicate, and validate the relevant skill overlap. This discussion leads to two hypotheses:

> **Hypothesis 1:** Skill-task match will vary substantially within occupations, majors, and expected match categories, especially among graduates outside common destinations.

> **Hypothesis 2:** Graduates whose transcript-derived skill portfolios are more closely aligned with the task demands of their jobs will earn more than otherwise similar graduates, even after accounting for the overall economic return to their skill portfolios, credentials, and occupations.

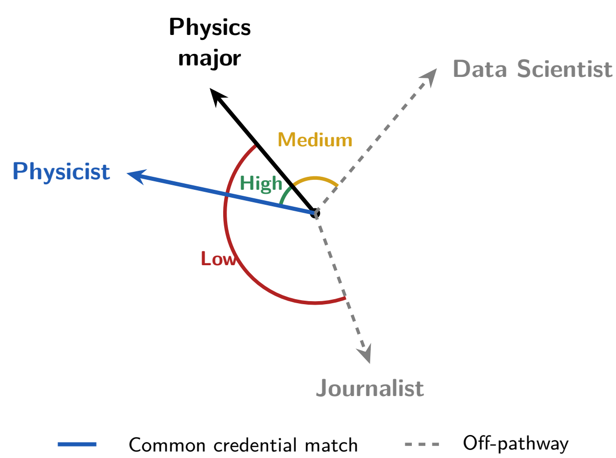

Figure 1: The "skill-task" space: Arrows depict the simulated "Physics Major" student (Black) and three target occupations as vectors emanating from a common origin. Line style encodes credential-to-occupation match (solid colored = common, dashed gray = off-pathway). Color of the arc and label at the origin encodes skill-task match (green = high, gold = mid, red = low).

3.3 Who gets to use their college skills? Disparities in matching and earnings gaps

If skill-task match is a central mechanism through which college-acquired skills are converted into labor-market rewards, who gets access to jobs that allow those skills to be used? I address this question by focusing on two dimensions of earnings inequality among BA holders: race and class background. Large racial and class-based earnings gaps persist in the United States. While a substantial share of these gaps is explained by differences in educational attainment, especially the unequal attainment of a BA degree, higher education does not eliminate inequality among degree holders. Racial earnings disparities persist among college graduates (Zhou and Pan, 2023), and graduates from lower-income families earn substantially less than graduates from more affluent families after college (Witteveen and Attewell, 2017b; Laurison and Friedman, 2024).

Existing research on post-BA inequality has focused primarily on unequal educational inputs, including differences in college selectivity, field of study, and skill acquisition. Black and lower-income students are less likely to attend selective institutions and are more likely to sort into lower-paying fields of study (Bowen and Bok, 1998; Davies and Guppy, 1997; Grodsky, 2007; Witteveen and Attewell, 2017b; Bleemer and Mehta, 2023). These differences in educational inputs matter for later earnings. But they leave open a further question: are groups that are already disadvantaged in access to high-value educational inputs also disadvantaged in the process through which education is converted into work?

Several processes could contribute to race- or class-based under-matching. First, discrimination in hiring could funnel minority graduates into jobs and occupations that are less well-suited for their skills. If all workers have a preference for working in "matched" occupations, Black and White workers may form a labor market queue in which the latter group are more regularly hired at their preferred jobs due to processes of racial and class-based discrimination (Bertrand and Mullainathan, 2004; Pager et al., 2009; Rivera, 2012) or homophily in job networks (Kanter, 1977; Kmec, 2007). Such demand-side exclusions could in turn spill over into differences in job search behaviors. In anticipation of discrimination by employers, minority groups conduct broader, less targeted job searches (Pager and Pedulla, 2015), which could in turn produce under-matching. Moreover, minority and disadvantaged job seekers may put more weight on job security and immediate income rather than the more intrinsic benefits (Johnson et al., 2012; Kalleberg and Marsden, 2013) considered by their more advantaged peers – which may include skill-job alignment.

Privileged workers may also use networks and strategic knowledge about elite firms and their selection processes to effectively obtain positions that align with their skills (Rivera, 2012). In particular, while human capital-intensive firms may be overwhelmingly White, minority groups may prefer employers that are "less White", even if those are the firms where they could attain greater alignment. In this way, employer-based differences could translate into minority under-match due to sorting across firms. While human capital-intensive firms are overwhelmingly located in wealthy, densely populated urban areas (Kogan et al., 2017), Black individuals are far more likely than their White counterparts to live in poorer neighborhoods. In their study of the class gap in earnings in the UK, Laurison and Friedman (Laurison and Friedman, 2016; Friedman and Laurison, 2019) found that 15 percent of the gap is attributable to geographic differences. By limiting access to high-skilled and appropriate jobs, long-standing residential segregation could further undermine skill-task match.

> **Hypothesis 3:** Non-White workers and workers from lower-class backgrounds are more likely to be under-matched.

4 Constructing skill-task match

The construction of skill-task match proceeds in three steps. First, I build an occupation-by-work-activity matrix from O*NET, representing each occupation as a bundle of task statements, and use a Singular Value Decomposition (SVD) to recover a lower-dimensional task space from this matrix. Second, I embed course descriptions and O*NET work-activity statements using a sentence embedding model, and use these embeddings to map each course onto the O*NET work activities to which it is semantically closest. Third, I aggregate course-level work-activity scores across each student's transcript to construct a student-level skill profile, and project both student skill profiles and occupation task profiles into the same SVD-reduced task space. Skill-task match can then be measured as the cosine similarity between these two profiles.

I begin by constructing the occupational task space. The starting point is an occupation–skill matrix $\mathbf{M}$, where rows index occupations ($m = 873$ at the SOC-detailed level) and columns index O*NET work activities ($n = 2,087$). Work activities are standardized descriptors of job tasks compiled by the U.S. Department of Labor, providing information on the work activities required across occupations (for example, "analyze data to identify trends," "develop statistical models," or "coordinate project activities"). Unlike broader skill or occupation categories, work activities reflect the specific tasks workers perform on the job, and therefore provide a comparable representation of task content across occupations. Entry $M_{op} = 1$ if work activity p is associated with occupation o in the O*NET activity statements.

This raw matrix represents each occupation as a bundle of work activities and could, in principle, be used directly to calculate distances between occupations. However, because this representation is high-dimensional and sparse, occupations that require similar forms of work may not share many identical work activity indicators. For example, physicists and data scientists may both require quantitative reasoning, modeling, and analysis, but these similarities may be distributed across different activity statements. To capture this higher-order similarity between tasks and occupations, I apply a Singular Value Decomposition,

$$\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\top},$$

and retain the first $k = 24$ dimensions, selected using an elbow criterion on the singular-value scree plot:

$$\mathbf{M}_k = \mathbf{U}_k\boldsymbol{\Sigma}_k\mathbf{V}_k^{\top}.$$

This decomposition embeds both occupations and work activities in a shared low-dimensional space. Occupations are represented by the rows of $\mathbf{U}_k\boldsymbol{\Sigma}_k$, while work activities are represented by the rows of $\mathbf{V}_k\boldsymbol{\Sigma}_k$. The resulting geometry captures higher-order task similarity: two occupations can be close not only because they share the same work activities, but because their work activities tend to co-occur with similar sets of other tasks across the occupational structure (see Han and Cheng, 2026). In this space, physicists and data scientists are more similar than physicists and journalists.

Table 2: Illustrative low-dimensional representation of the skill-task space. The first row gives a transcript-derived skill vector for a physics major, while subsequent rows give occupation task vectors based on O*NET work activities projected into a simplified three-dimensional SVD space. The final column reports the cosine similarity between the physics-major vector and each occupation vector. Entries are shown as binary 1/0 indicators for illustrative purposes only.

	Scientific / modeling domain	Applied data / business domain	Interviewing / writing domain	Skill-task match
Physics major	1	0	0	–
Physicist	1	0	0	1.00
Data Scientist	1	1	0	0.71
Journalist	0	0	1	0.00

The second step is to map each college course to its most semantically related work activities. To do this, I first embed each cleaned course paragraph (see next section for details) as well as the O*NET work activity statements into a vector space using a sentence embedding model. For each (course, work activity) pair (c, p), I compute the cosine similarity between the course paragraph embedding $\mathbf{e}_c$ and the work activity descriptor embedding $\mathbf{s}_p$, resulting in a course–work activity similarity matrix $\mathbf{A}$ with approximately 290,000 rows (courses) and columns for each O*NET work activity. Each entry a_{cp} represents the semantic proximity between a course description and a specific occupational activity, ranging from -1 (opposite meaning) to $+1$ (identical meaning). Because each course is plausibly unrelated to most work activities, for each course c, I retain only the top 5% of work activities by cosine similarity to the course description (the $q = .95$ threshold), setting all other similarities to 0. The results are not sensitive to the exact choice of threshold.

Finally, I compute a student's skill profile via the simple sum of each work activity score across courses that the student takes. In alternative specifications, I weight each course by a student-specific term reflecting course intensity and performance, and apply an element-wise concave transformation which models diminishing marginal returns to repeated exposure to similar content. I also estimate a specification that controls for course dissimilarity, defined as the average pairwise dissimilarity across all courses a student takes, to test whether my measure of match may partly capture differences in curricular breadth. The results are not sensitive to these alternative modeling procedures (see Appendix D). I project each person's skill profile as well as each task profile into the same 24-dimensional task space obtained via the Singular Value Decomposition. I use the resultant individual-level "Skill Profile" as a control in all the analyses that follow, in order to adjust for differences in detailed course investments even net of field of study.

Having constructed a course-derived skill vector for each student and an occupation-derived task vector for each job (in the same 24-dimensional task space), I measure skill-task match D_i as the cosine similarity between the two:

$$D_i := \text{sim}(Z_{1i}, Z_{2i}) = \cos\theta = \frac{Z_{1i} \cdot Z_{2i}}{||Z_{1i}|| ||Z_{2i}||}, \tag{1}$$

This score, which lies between -1 and 1, captures the extent to which a graduate's actual skill investments align with the tasks required in their job. A higher cosine similarity between these vectors indicates a stronger match between an individual's educational background and their job's tasks, while lower similarity reflects skill mismatch. This procedure is shown graphically in Figure 2, and illustrated for a physics major and three example occupations in Table 2 using a simplified three-dimensional version of the SVD-reduced task space.[3] Here, the physics profile overlaps most closely with the task vector for physicists, somewhat less with the task vector for data scientists, and least with the task vector for journalists. The resulting cosine similarities show how the same educational profile can be strongly aligned with a common occupational destination, partially aligned with an uncommon but task-adjacent occupation, and weakly aligned with an occupation whose work activities draw on a different set of skills. I use the similarity scores estimated in Equation 1 as both an outcome and a treatment measure: as an outcome to examine variation in match as well as inequalities in match, and as a treatment to assess the economic effects of match.

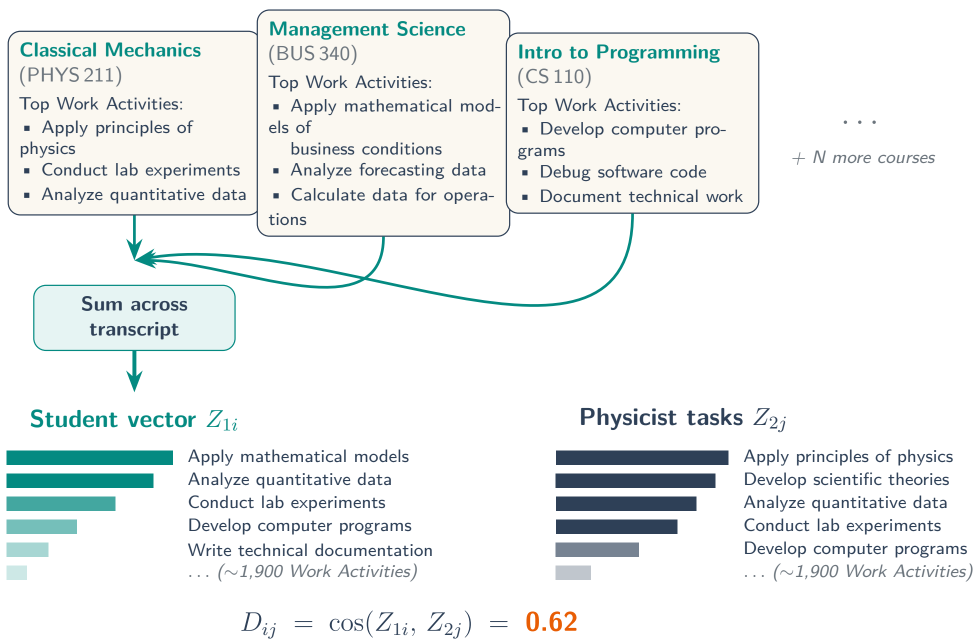

Figure 2: The "skill-task match" framework: each student's transcript is transformed into a continuous, multidimensional skill profile by aggregating embedded course descriptions in the O*NET work activity task space; each post-college occupation is represented in the same task space using its O*NET work activities; and skill-task match is the cosine similarity between the two vectors. This approach captures forms of alignment between schooling and work that extend beyond binary credential-to-occupation linkages.

[3]To aid intuition, Table 2 reports entries as binary 1/0 indicators; in the actual implementation, entries are continuous SVD coordinates (real-valued loadings on each latent task dimension), and the cosine similarity is computed from those continuous vectors.

5 Data

I analyze skill-task match among BA graduates in the context of one U.S. state's public 4-year college sector using a new anonymized dataset that links several sources of administrative data from that state.[4]

The first dataset is earnings data. These data come from the state's Unemployment Insurance (UI) program, which requires nearly all employers to submit quarterly reports on employee wages, hours, and occupations. These reports cover most private, local, and state public-sector workers. Notably, several groups—self-employed, federal employees, and active-duty military—are not covered. My main outcome is annualized earnings, measured in financial years 2023 and 2024, and adjusted to 2024 dollars using the CPI-U. Annual earnings are constructed by summing observed quarterly earnings, and workers with multiple jobs are assigned to the employer where they received the largest share of their compensation, provided they are observed at their primary employer for at least two quarters in any given financial year.[5]

Employers report occupation codes for each employee as part of their UI quarterly filings, using the standard six-digit Standard Occupational Classification (SOC). Coverage is partial: a non-trivial share of worker-quarter observations carry no SOC code, primarily because some employers either do not report occupation or report invalid codes. To assess whether this missingness is consequential, Appendix C compares the distribution of demographic covariates, prior academic preparation, and earnings between workers with and without SOC codes. Overall, the two groups are substantively similar on observables, though with some notable differences, particularly regarding employment: individuals with a reported SOC code work at larger employers that are more likely to be multi-establishment companies, and earn more. This pattern suggests that SOC-code availability reflects employer reporting capacity as well as worker characteristics: larger and more formalized firms may be better equipped to comply with mandatory occupation-reporting requirements.

Educational records come from the state's postsecondary system as well as the high school system. Postsecondary files provide information on institution attended, GPA, credits, degree attainment, and postsecondary transcripts on courses taken in college. To ensure that education is complete before analyzing wages, I assign to each individual their highest credential attained in the *prior* financial year. If a student obtains multiple BA credentials in any given year, I attach to that student an in-state public degree if obtained, followed by a public or private 4-year out-of-state degree. If a student obtains multiple in-state BAs in a given year, I follow a rank ordering in which

[4]These datasets were linked to each other at the individual level by the state data agency, using social security number and identifying information.

[5]I drop worker-firm observations with missing information on these attributes.

I assign students to the most selective degree obtained. Students who pursue graduate education are retained in the sample, but their degree attainment is coded to their highest BA institution. I also use enrollment records to identify a student's credits, and cumulative GPA. High school transcript records provide GPA, school attended, and demographic characteristics (e.g., gender, race/ethnicity, disability status, and free and reduced-price lunch [FRPL] eligibility), as well as information on SAT and ACT scores provided by the College Board.

Finally, I develop skill representations of each course taken by students in the sample using publicly available course catalogs for each college-year in my sample. I obtain course catalogs between 2005 and 2022 for each of the five colleges via a combination of web-scraping (for colleges providing online catalogs) and Optical Character Recognition (OCR) procedures (for colleges providing PDF versions of their catalogs). Each course's catalog text was pre-processed (using the procedure described in *Processing course catalog text* below) and then put into an embedding space yielding a numeric representation for each course. I then matched each course, as it appears in the administrative transcript data, to the numeric representation of the course as obtained from this embedding procedure. Because course names in the online catalogs differ slightly from course names in administrative transcript data, each catalog course was linked to the corresponding course in the administrative records via a combination of fuzzy string matching and hand-coding, restricting candidate matches to courses with the same CIP code and course number within the appropriate academic year. Figure A.1 in Appendix A shows that I successfully match 82% of all 354,806 unique courses taken by students over the period to a course catalog entry. I construct individual skill profiles by collapsing over the numeric representations of each course they took, as described in the previous section.

Processing course catalog text. I assemble course paragraph descriptions from these course catalogs using a multi-step procedure. Note that this procedure was performed on the public course catalog text; only the embedding representations of each cleaned catalog paragraph were matched to course names in the administrative records as described above. I first segment each course catalog description into sentences. Across the five institutions, 253,383 course descriptions with valid substantive text yield 741,961 sentences (that is, ≈ 2.9 sentences per description). A key implementation challenge was in accurately removing administrative and logistical content from course catalog descriptions. In preliminary tests, I found that boilerplate sentences (e.g., references to course prerequisites, scheduling, and assessment policies) artificially inflate descriptions' similarity to education- and pedagogy-related work activities, biasing match measures upward for any course with a typical syllabus footer. I therefore filter sentences in two stages. First, a rule-based keyword filter discards sentences containing administrative or logistical language; this removes approximately 28% of sentences overall, and 6% of courses. Second, surviving sentences are submitted to an LLM for binary classification of skill relevance; only sentences classified as skill-relevant are retained (see Appendix A for details).[6] Sentences passing both filters are concatenated in original order into a single paragraph per course; courses with fewer than five surviving words are dropped, and course-CIP descriptions are appended to provide additional content. The final output is one cleaned paragraph per course. Figure 3 illustrates this cleaning process for a representative catalog entry. Overall, 82% of courses are successfully matched to the corresponding catalog entry; a further 11% of courses are lost because the keyword and LLM filters deem that the course catalog description contains no substantive (skill-related) content. Counts of courses remaining at each filter stage are reported in Appendix Figure A.1. On average, 77% of a student's transcript courses are successfully matched and embedded after the keyword and LLM filters (median 81%; bottom decile below 57%). Although I successfully measure the majority of course investments for most students, it remains possible that the unmeasured courses reflect skill investments that are aligned or unaligned with their work tasks, which could bias my measure of match.[7] To test for this possibility, I repeat the main analyses on the subsample of students for whom I am able to embed at least 80% of courses; the substantive results are unchanged (Figure D.1 in Appendix D).

I also construct common credential-occupation match measures as a point of comparison with my text-based measure of match. To do this, I draw on the ACS 5-year sample 2019–2023. For each field-education-level cell, I identify matched occupations as follows. I first select the ten largest SOC 3-digit minor groups by weighted employment count among workers from that field and education level. I then compute an actual-to-expected (A/E) ratio for each, defined as observed employment divided by the employment that would be expected if graduates from that field were distributed across occupations in the same proportions as all workers at that education level — that is, $\text{A/E}_{fgl} = n_{fgl} \,/\, (n_{gl} \cdot n_{fl}/n_l)$, following Bol et al. (2019). The two occupations with the highest A/E ratio among the top ten constitute the matched occupations for that cell. An individual is coded as matched (= 1) if they work in either of the two matched SOC minor groups for their field and education level, and unmatched (= 0) otherwise.

I focus my analyses on BA recipients who began their postsecondary careers at one of five (out of seven) public four-year colleges in the state, graduated with a BA between 2008 and 2022, and are observed with positive UI-covered earnings in the relevant period. This restriction is motivated by measurement feasibility. Constructing skill-task match requires a comprehensive transcript-derived skill profile, which in turn

[6] Note that the LLM was run only on course catalog text; no individual-level data was passed through any LLM.

[7] For example, I do not see the courses that students take in institutions outside the public state system.

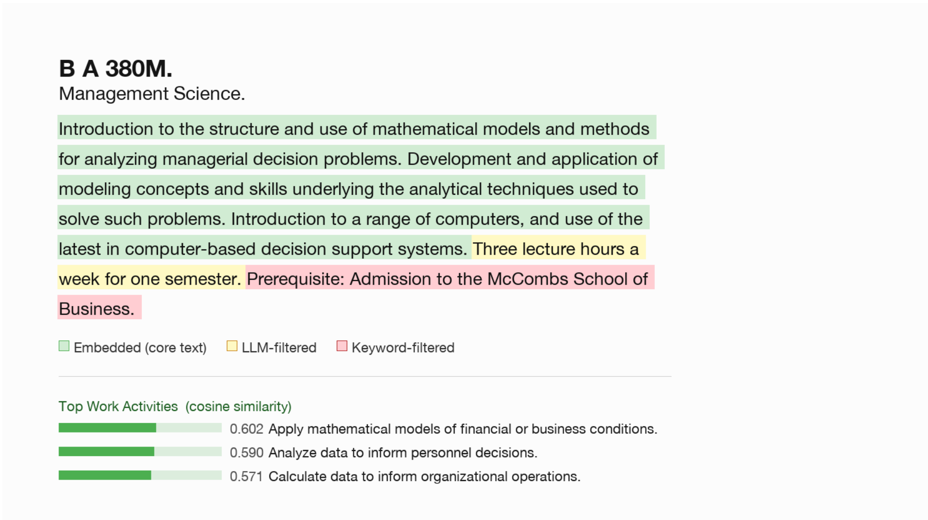

Figure 3: Illustration of the catalog-cleaning pipeline for a representative course. Raw catalog text is segmented into sentences; sentences containing administrative or logistical content are dropped by a rule-based keyword filter; remaining sentences are submitted to GPT-4o-mini for binary skill-relevance classification; surviving sentences are concatenated into a single cleaned paragraph used as input to the SBERT embedding step.

requires linking transcript courses to course-catalog text from the institutions students attended. Students who begin at two-year institutions and later transfer to four-year colleges typically attend a larger and more heterogeneous set of institutions, which would substantially expand the number of course catalogs that must be collected, processed, and linked. I therefore focus on four-year entrants in this analysis.[8] Extending the analysis to two-year college entrants, associate-degree holders and certificate programs is an important avenue for future work.

Table 3 presents descriptive statistics for the core sample. Of the 74,811 workers in the overall sample of eligible BA graduates, 69% are observed with in-state earnings over the 2023-2024 period. Of these workers, 71% have employer-reported SOC codes. Workers with missing SOC codes for their primary employer-year observation are dropped from the matched-occupation analyses, since my measure of skill-task match requires both a course-derived skill profile and an occupation-derived task profile. Among workers with valid SOC codes, approximately 3,000 are excluded from the main analysis because their reported occupations do not have corresponding O*NET

[8] Students who transfer between four-year institutions are retained in the analytic sample. I do not explicitly drop BA graduates who also obtained associate degrees or other certificates, but in practice, students who begin their postsecondary studies at a 4-year college are unlikely to attain these credentials.

task ratings, which are required to construct occupation task profiles. This leaves me with a final analytic sample of $N = 32,694$.

Table 3: Sample descriptives for BA completers.

Characteristic	Mean / proportion
Female	0.567
Black	0.029
Hispanic	0.070
White	0.667
Asian	0.194
FRPL (ever)	0.222
Final GPA	3.54
Has earnings	0.691
Has wage	0.690
Employer reported SOC code	0.491
Employer reported SOC code (cond. on earnings)	0.710
Log earnings	11.288
Log wage	3.898
N	74,811

5.1 Validating the measure

I validate the measure in two complementary ways. First, I examine whether skill-task match is higher along common credential-to-occupation pathways, and whether skill-task match varies in ways we would expect across different fields of study (see Appendix B). Table B.2 shows that skill-task match is positively correlated with credential-based match ($r = .563$), indicating that graduates who enter conventional occupational pathways for their field also tend to have higher skill-task match scores. Consistent with this pattern, graduates classified as matched under the credential-based measure have substantially higher average skill-task match scores than those classified as unmatched (.609 versus .267). Figure B.1 further shows that average match scores vary in expected ways across fields of study. Fields conventionally characterized by weaker school-to-work linkages, such as the liberal arts, humanities, and social sciences (Bol et al., 2019), have the lowest average match scores, while fields with more structured pathways into work, such as health, engineering, and business, have the highest. These patterns provide an important validity check: although the skill-task measure is not defined in terms of common pathways, the measure recovers broad field-level differences that prior credential-based approaches would lead us to expect.

Second, I show that the measure responds correctly to the underlying course composition of a student's profile. Specifically, I simulate physics-major students whose 30-course transcripts contain $N \in \{0, \ldots, 30\}$ physics courses (randomly drawn from the corresponding portions of the course catalog) plus $30 - N$ non-STEM courses. For each N, I project the simulated transcript into the latent skill space and compute its cosine similarity to three target occupations: Physicists, Data Scientists, and Journalists. Figure 4 shows the result. As expected, similarity to Physicists rises sharply with N. Crucially, similarity to Data Scientists also rises with N—capturing the skill transferability the measure is designed to recover, since physics coursework develops mathematical and computational tasks that data-scientist work requires—while similarity to Journalists declines.

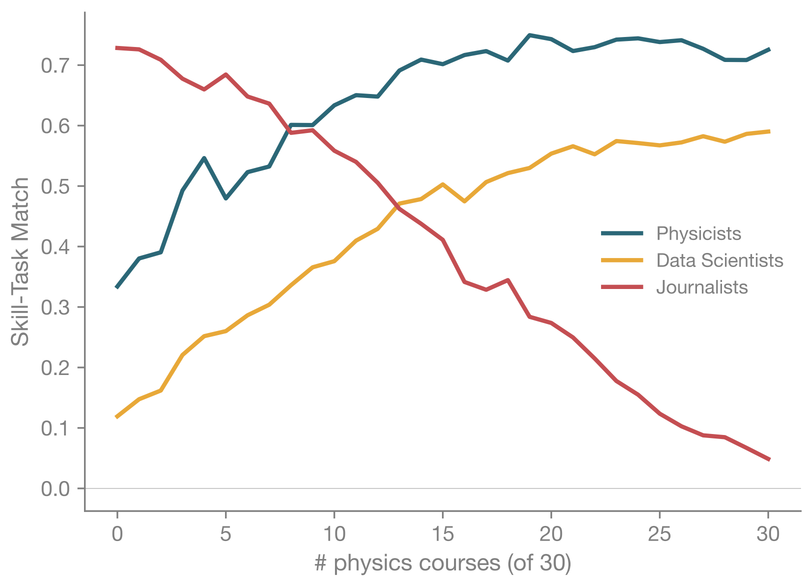

Figure 4: Skill-task match between simulated physics-major transcripts and three target occupations (Physicists, Data Scientists, Journalists), as a function of the number of physics courses (out of 30) on the transcript. Similarity is computed in the work-activity-based latent skill space.

6 Results

6.1 Skill-task match: variation and variance decomposition

I begin by characterizing the variation in skill-task match among individuals in my sample. Table B.1 shows that the average match score is .355; match scores have a standard deviation of .266, a 25th percentile of .125, and a 75th percentile of .585. We can gain an economic interpretation for these numbers in two ways. First, we can consider fixing a person's occupation, and asking how much more coursework that person would need to take in a field associated with the occupation to obtain a one-standard-deviation increase in match. The validation simulation reported in Figure 4 shows that, for a fixed occupation, a one-standard-deviation increase in skill-task match corresponds to replacing approximately one third of an average student's coursework with field-aligned courses, and moving a student from the 25th to the 75th percentile of match corresponds to replacing approximately 80% of their coursework. Second, we can consider fixing a person's coursework, and asking what sort of occupations would correspond to a change in match score. To give an illustration of this, for business majors who take more math-based courses, a one-standard-deviation increase in match corresponds to the difference between a math-oriented business major who becomes a data scientist, and a math-oriented business major who becomes a credit analyst.

What drives this variation? I examine the extent to which skill-task matching D_i varies within majors, occupations, schools, and a measure of making a common credential-to-occupation match. Specifically, I implement a variance decomposition of match as follows:

$$\text{Var}(D_i) = \underbrace{\text{Var}(\mathbb{E}[D_i|C_i])}_{\text{between-}C_i} + \underbrace{\mathbb{E}[\text{Var}(D_i|C_i)]}_{\text{within-}C_i},$$

where C_i denotes majors, occupations, or following a common match pathway. This decomposition allows me to quantify how much of the overall variation in skill-task congruence arises from systematic differences between categories (e.g., certain majors being better matched than others) versus within categories (e.g., students with the same major experiencing very different match outcomes).

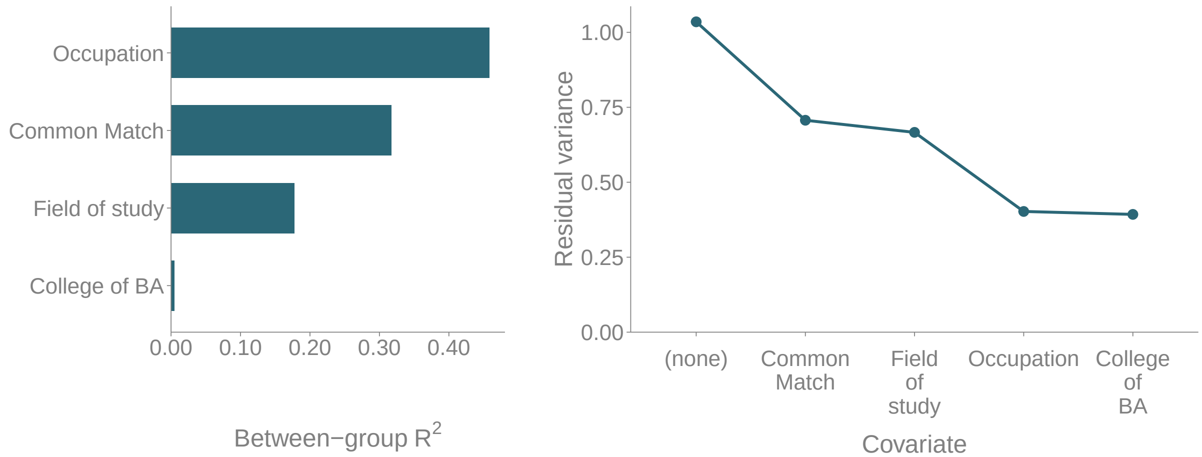

Figure 5: Left panel: variance decomposition of the skill-task match score. Bars show the share of total variance attributable to (i) field of study, (ii) occupation, (iii) BA college, and (iv) following a common credential-to-occupation pathway, each entered separately as a categorical predictor. Right panel: residual variance for skill-task match after the same predictors are added sequentially in a nested model.

The left panel of Figure 5 reports the results of this exercise. BA college explains the lowest amount of variation: almost all of the variation in skill-task matching occurs among graduates of the same college, implying that specific colleges do not systematically sort their graduates into better-matched jobs than others. Field of study explains a modest portion of the total variation in skill-task matching, at 18%. Occupation explains the largest portion, at 46%, but a full half of the variation in match occurs within occupation groups: among individuals obtaining the same occupation, there are dramatic differences in the extent to which individuals are using their college-acquired skills. Importantly, common pathways between occupations and credentials contribute under one-third of the total variation in matching.

One limitation of the preceding variance decomposition is that it considers each factor independently. Yet, some of these factors are likely to be correlated: field of study, for example, partly shapes the occupations graduates enter, and one's degree-granting institution shapes the fields of study available. To assess how much explanatory power each factor contributes net of the other factors, I estimate a series of nested regressions where I sequentially add common credential-occupation pathway, field of study, occupation, and BA college to a regression model of skill-task match. The results are shown in the right panel of Figure 5. The results are consistent with the variance decomposition exercise. Just under one third of the variation in skill-task match reflects whether an individual follows a common or uncommon pathway. Adding field of study to the regression model only decreases the residual variance by a small amount, whereas adding occupation decreases the variation considerably. In other words, some occupations tend to have systematically higher levels of alignment than others. Finally, adding BA college to the model does not explain additional variation in skill-task match. Collectively, field of study, occupation, BA college, and common match pathways account for 62% of the total variance in skill-task match. While the skill-task match measure therefore partly overlaps with these measures, skill-task match captures dimensions of the school-to-work transition that are independent of common credential-to-occupation linkages, providing support for Hypothesis 1.

6.2 Returns to skill-task match

Is skill-task match associated with earnings? To provide a descriptive test of this proposition, Figure 6 shows a binned scatterplot of log earnings against standardized skill-task match score, separately for graduates whose job is on a common credential-to-occupation pathway for their major and graduates whose job is off-pathway. The size of each bin is proportional to the number of graduates in each. I annotate several of the bins with the modal occupation in each.

The figure shows substantial variation in match, across both common and uncommon pathways. Examining just variation along the x-axis of each panel shows that, among workers who do and do not make common matches, there is a wide range in who leverages their college-acquired skills on the job. This reflects the result in Figure 5, that most of the variation in skill-task match is within, rather than between, common-match categories. Further, this variation in match is associated with earnings. Individuals with higher scores of skill-task match generally have higher average earnings than those with lower scores.

Interestingly, the shape of the association between skill-task match and earnings is similar both for individuals who follow a common pathway and those who do not, steepest for those with the lowest match scores, and plateauing for those with the highest scores. The steeper gradient between matching and earnings among those with lower match scores may partly reflect occupation attainment: individuals with the lowest match scores are more likely to work in jobs involving routine sales, administrative, and clerical tasks – jobs which are well-known to often pay less. Occupational sorting also appears to explain some of the outlier points in the figure. Education majors working in the teaching profession are among the best matched workers following a common pathway, but are also relatively low paid. Business and social science majors working as financial analysts are poorly skill-task matched, but highly paid. At the same time, skill-task matching does not appear to be coterminous with occupation-based sorting. Individuals working in several generally high-paying occupations, but with lower skill-task match scores, earn less than those in the same occupations but with higher skill-task match scores. Nevertheless, to ensure I am not simply picking up an occupational sorting effect, in the earnings models presented below I include occupation fixed effects. This allows me to ask how skill-task alignment is associated with earnings among BA holders working in the same occupation, but who vary in their skill-job alignment.

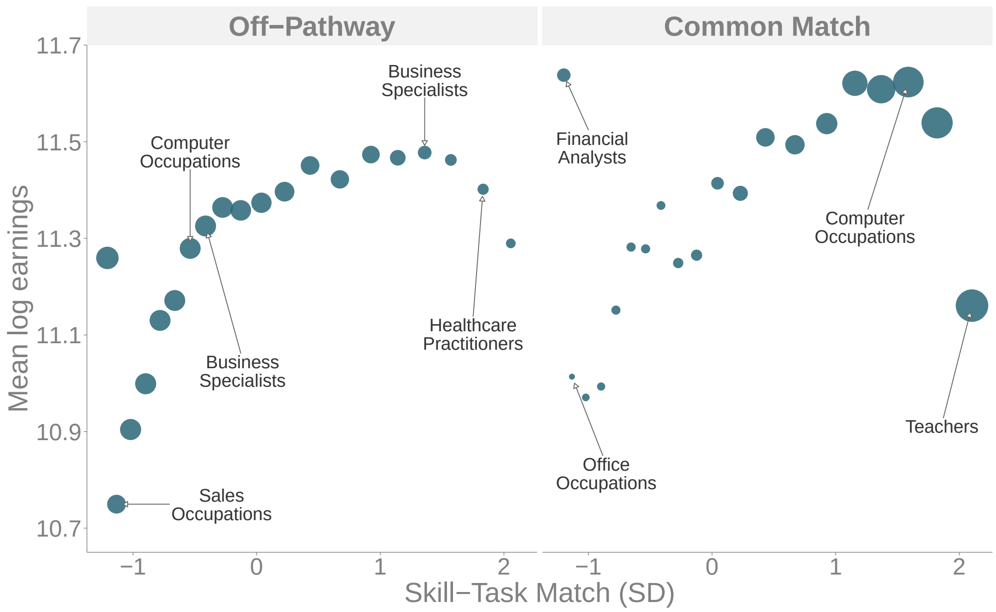

Figure 6: Binned scatterplot of log earnings on skill-task match score, separately for graduates on a common credential-to-occupation pathway and off the common pathway. Point size is proportional to the number of graduates in each bin.

To be sure, a large portion of this earnings premium may be driven by the characteristics of graduates who successfully make a match. Appendix Table B.3 shows that highly matched candidates are more likely to be White, have higher GPA and standardized test scores, and are less likely to have been on free and reduced-price lunches (FRPL) in high school. To isolate these characteristics from the role of match, I estimate the partial effect of match on earnings, conditional on skills and other covariates. Specifically, I fit the following model:

$$Y_i = \beta_1 D_i + \beta_2 X_i + \beta_3 Z_i + \alpha_i + \epsilon_i, \tag{2}$$

where Y_i is log annualized earnings, X_i denotes pre-college and college covariates (high-school GPA, SAT score, high school fixed effects, field of study, degree-granting institution, race, gender, free and reduced-price lunch status), and α_i are occupation fixed effects (6-digit SOC codes). I also control for Z_i, which denotes the student's skill-embedding vector—a low-dimensional control for the student's college skill investments. I estimate parallel specifications that replace D_i with an indicator for making a common credential-to-occupation match, as well as specifications that include both skill-task match and common match jointly.

Figure 7 displays estimates of the earnings return to matching across each of these three specifications. The coefficients in blue show the estimated returns to common match (top row) and skill-task match (bottom row) separately. When estimated in separate models, both skill-task matching and matching along a common credential-to-occupation pathway yield a sizeable earnings return. In particular, a one-standard-deviation increase in skill-task match is associated with an earnings return of approximately 7%. Because of the rich set of controls, the skill-task match effect can be interpreted as the earnings difference between two BA holders with the same GPA, field of study, and from the same college who work in jobs with identical average earnings overall, but one of whom works in a job that is better aligned with their college skills.[9] The earnings gain to skill-task match is also comparable in size to social skill complementarities reported by Deming (2017), who showed that complementarity between a worker's social skills and social skill job requirements increases earnings by about 5%.

How distinct is the earnings return to skill-task match from the return to making a common pathway match? To address this, the coefficients in yellow present estimates of match earnings effects from a model which estimates returns to common match and skill-task match when both measures are placed in a model together. When both measures enter together, the estimated return to skill-task matching stays constant at 7%; by contrast, the estimated return to common matching attenuates to zero. That is to say, credential-to-occupation pathways do not carry an earnings premium over and above the actual alignment between graduates' college-acquired skills and the tasks of their jobs. Instead, the premium associated with common pathway match appears to arise largely because these pathways are more likely to place graduates in work that uses the skills their education developed. Together, these results provide support for Hypothesis 2.

Table 4 reports the corresponding regression estimates across increasingly restrictive specifications. Without covariate adjustment, a one-standard-deviation increase in skill-task match is associated with a 0.139 point increase in log earnings. This association only declines slightly when pre-college controls are included, implying that

[9] Notably, my estimated payoff to matching along a credential-to-occupation pathway is smaller than the reported estimate in Bol et al. (2019). This attenuation is almost entirely driven by the inclusion of occupation fixed effects in my specification, which Bol et al. (2019) do not include in their models.

pre-college achievement and resources are not strongly predictive of making a match. By contrast, including college controls attenuates the return to skill-task match by 26%, which primarily reflects the association between field of study and match as shown in Figure 5. Including occupation fixed effects attenuates the estimated match effect by another 30%, implying that part of the association between earnings and alignment comes from the fact that the most misaligned workers are found in jobs that are themselves low-paying (a point reflected in Figure 6). Finally, in Appendix E, I examine whether the earnings return to individual skill-task match reflects major-level alignment rather than student-specific course-taking, by comparing individual skill-task match to major-task match: the alignment between a graduate's job tasks and the average skill profile implied by their field of study. The results show that the return to skill-task match is not reducible to the average skill profile associated with a graduate's major, implying that students' specific course trajectories capture worker–job fit beyond what can be inferred from field of study alone.

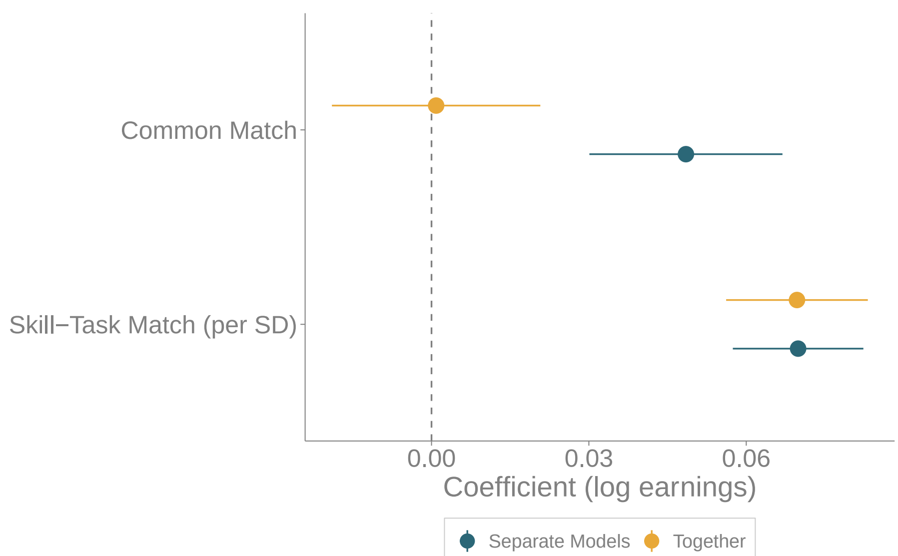

Figure 7: Earnings returns to skill-task match vs. common credential-to-occupation match across three model specifications. The key predictor is a one-standard-deviation increase in skill-task match (or a binary indicator of making a common match). Models control for pre-college characteristics (high-school GPA, SAT/ACT scores, race/ethnicity, gender, free and reduced-price lunch status), BA college fixed effects, field-of-study fixed effects, high-school fixed effects, occupation fixed effects (6-digit SOC codes), the low-dimensional transcript-derived skill profile, and the common credential-to-occupation match indicator. Point estimates with 95% confidence intervals.

Table 4: Earnings returns to Skill-Task Match across nested control specifications.

	M1	M2	M3	M4	M5	M6
Skill-Task Match	0.139	0.134	0.103	0.102	0.070	0.070
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)
Common Match	–	–	–	–	–	0.001
						(0.010)
Pre-college controls		✓	✓	✓	✓	✓
College / Field of Study controls			✓	✓	✓	✓
High-school fixed effects				✓	✓	✓
Occupation fixed effects					✓	✓
Common Match						✓

Notes: Standard errors in parentheses. **Pre-college controls** comprise student demographics (female, Black, Hispanic), GED status, high-school graduation year, indicators for ever experiencing homelessness, free/reduced-price lunch, special education, or English-language-learner status, high-school GPA, and the standardized SAT/ACT composite score. **College / Field of Study controls** comprise BA-college fixed effects, field of study, college GPA, and the student's course-profile embedding (capturing the average skill content of courses taken). **Occupation fixed effects** are 6-digit SOC codes. **Common Match** is an indicator equal to one if the student follows a common credential-to-occupation pathway and zero otherwise.

6.3 Race- and class-based gaps in match

Finally, I assess the role of skill-task match in driving demographic earnings gaps among college graduates. To do this, I first regress measures of match on three demographic characteristics: Hispanic (versus White), Black (versus White), and an indicator of class background (whether an individual ever received free and reduced-price lunches in high school). In each specification, I include narrow controls for high school GPA, BA institution, field of study, as well as workers' skill embeddings. These gaps ask: among students with the same pre-college academic preparation, the same college and field of study, and the same latent skill profile, are minority or low-income students less likely to make a match? Similarly to the previous section, I conduct this exercise separately for skill-task matching as well as common pathway matching.

Figure 8 presents the results. The coefficient estimates in blue show estimated gaps in skill-task match, while the yellow point estimates show estimated gaps in making a common match, conditional on pre-college covariates, BA college and field of study, and individuals' skill profiles. Inequalities appear primarily off, rather than on, common match pathways. Demographic gaps in making a common match are small and (with the exception of the Black-White gap in match) statistically indistinguishable from zero. By contrast, under a skill-task match measure, the same comparisons recover large negative gaps for Black graduates, Hispanic graduates, and ever-FRPL graduates. The gap is particularly large between Black and White graduates, corresponding to roughly a .19 SD gap in match. While these results alone do not provide evidence of mechanisms, they do provide some suggestive evidence that explicitly racialized mechanisms (such as differences in job search behaviors and hiring discrimination), rather than mechanisms due to financial resources alone (such as the importance of bridging funds), are most important in driving gaps in matching outcomes.[10] [11]

While Figure 8 shows that minority and low-income students are less likely to obtain jobs that leverage their college skills, this does not itself show how much these gaps matter for earnings gaps between these groups. To quantify the importance of these gaps for earnings disparities, I translate these match gaps into the earnings domain via a Kitagawa–Oaxaca–Blinder decomposition. Specifically, I decompose the total disparity in earnings between racial and family-income groups (among college graduates) into five components: pre-college characteristics, BA college, field of study and skill profiles, occupations and skill-task matching. Specifically, let X_i denote the column vector

[10] The caveat here is that FRPL-status could act as too crude a proxy for family income. A continuous measure of family income could explain more of the Black-White gap in match, but such a measure was unavailable for my study.

[11] My finding of a small Black disadvantage in obtaining a common credential-to-occupation match aligns with that in Lu et al. (2025), who also find a small Black disadvantage in this measure of match using the Survey of Income and Program Participation.

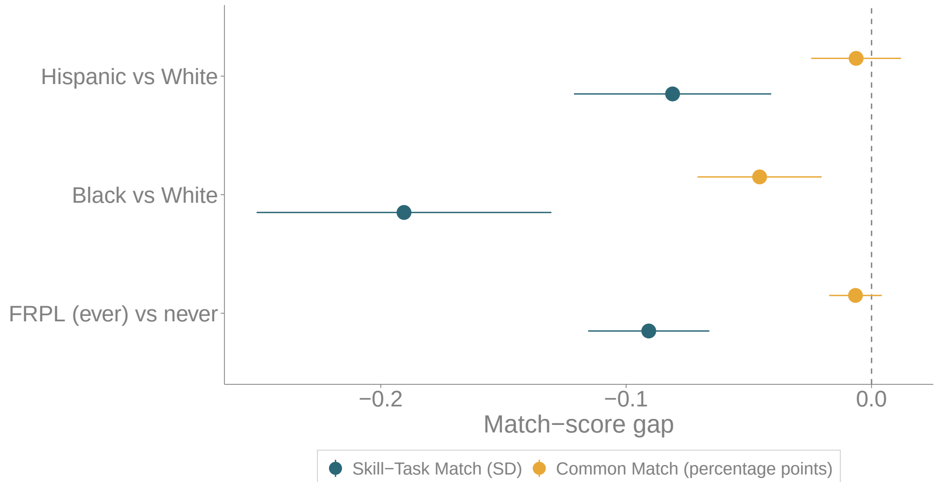

Figure 8: Demographic gaps in skill-task match versus common credential-to-occupation matching. Each panel reports the within-group difference (Black–White, Hispanic–White, FRPL-ever vs. never-FRPL) in mean skill-task match (blue) and in the probability of common match (yellow), conditional on pre-college covariates, college-level controls, field of study, and skill embeddings. 95% confidence intervals shown.

$X_i = [X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}]'$, where X_{i1} denotes *pre-college characteristics* (high school GPA, English as a second language status, SAT/ACT score, and FRPL status for the racial earnings gaps), X_{i2} denotes *BA college*, X_{i3} denotes *field of study and skill embeddings*, X_{i4} denotes *occupations*, and X_{i5} denotes *skill-task match*. Then, the total disparity between groups can be decomposed as (Blinder 1973; Oaxaca 1973)

$$\widehat{\Delta} = \underbrace{\sum_{j=1}^{5} \hat{\beta}_j^B \left(\bar{X}_{ij}^W - \bar{X}_{ij}^B\right)}_{\widehat{\Delta}_{\text{Explained}}} + \underbrace{(\hat{\alpha}^W - \hat{\alpha}^B) + \sum_{j=1}^{5} \left(\hat{\beta}_j^W - \hat{\beta}_j^B\right) \bar{X}_{ij}^W}_{\widehat{\Delta}_{\text{Unexplained}}} . \tag{3}$$

where $\hat{\beta}_j^B$ and $\hat{\beta}_j^W$ denote the slope coefficients on X_{ij} for minority (e.g., Black) and majority (e.g., White) students, respectively, and $\bar{X}_{ij}^B$ and $\bar{X}_{ij}^W$ denote the corresponding group-mean values of X_{ij}. In this decomposition, $\widehat{\Delta}_{\text{Explained}}$ consists of five components, each representing the contribution of each group of covariates described above. By contrast, $\widehat{\Delta}_{\text{Unexplained}}$ captures group differences in the "returns" to each set of covariates (Jann, 2008).

Figure 9 shows the results of this exercise. For each demographic comparison, the point estimates decompose the raw log-earnings gap (red bars) into the portions explained by each component. Each component captures the direct contribution of the component, net of the subsequent components included in the decomposition. For example, the "Pre-College" component captures the direct contribution of pre-college resources, net of BA college, Field of Study / Skill Profiles, Occupation, and Skill-Task Match.

Pre-college characteristics explain a small (and in all cases, statistically insignificant) share of the total disparity between groups. This reflects the fact that, once BA college, field of study / skill profiles, occupation, and skill-task match are accounted for, race- and class-based differences in pre-college resources hold no additional explanatory power for the total earnings differences between these groups. For the FRPL-non-FRPL gap, BA college explains a slightly positive portion of the gap, whereas for the racial gaps in earnings, it explains a negative portion of the total gap. This results from the fact that, conditional on pre-college academic performance, Black and Hispanic students are in fact *more* likely to attend the Flagship school than their White peers—meaning that college-going actually compresses the gap between racial groups. Together, field of study and course profiles explain an important portion of the gaps (32% of the Black–White gap, 13% of the Hispanic–White gap, and 16% of the FRPL-non-FRPL gap). Occupational sorting also explains an important component of each gap (although its contribution is less precisely estimated for the Black–White gap) reflecting previous findings that racial minorities are disproportionately sorted into lower-paying occupations and are less likely to attain positions of authority and managerial power (Tomaskovic-Devey et al., 2006; McTague et al., 2009; Kalev et al., 2006; Skaggs, 2009).

Importantly, skill-task match accounts for a non-trivial share of every gap (Hypothesis 3). For the Black–White gap, its contribution is on par with that of field of study and skill investments: skill-task match accounts for 25% of the Black–White gap in earnings among college graduates. The portion explained by match is smaller, but still non-trivial, for the Hispanic–White and FRPL-non-FRPL earnings gaps, capturing 5% of the Hispanic–White gap and 8% of the FRPL-non-FRPL gap. Importantly, the portion explained by skill-task alignment is *net of* occupation, which implies that race- and class-based sorting into skill-aligned work is an inequality-generating mechanism distinct from simply sorting into higher-paying work.

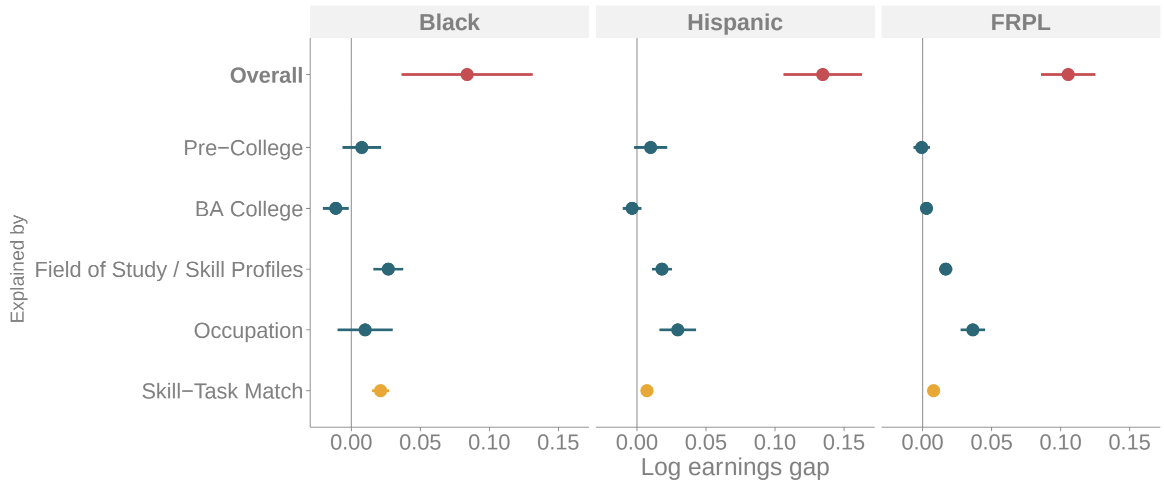

Figure 9: Kitagawa–Oaxaca–Blinder decomposition of the log-earnings gap, by group (Black–White, Hispanic–White, FRPL vs. never-FRPL). Within each panel, the top point (red) reports the unconditional log-earnings gap; subsequent points (blue) report the share of the gap attributable to covariate blocks (pre-college characteristics, BA college, field of study and course profiles, occupation); the bottom bar (gold) reports the share attributable to skill-task match. Each block's contribution is incremental over the preceding blocks. Skill profiles are the low-dimensional work-activity representation of students' transcripts. 95% confidence intervals from $B = 200$ bootstrap replicates.

7 Discussion

The transition from college to work is often understood in terms of formal categories: credentials, occupations, and the pathways that link them. This paper shows that these categories miss an important dimension of labor-market alignment: the transfer of college-acquired skills across occupational boundaries. I measure this alignment directly by developing a novel approach to characterizing the skill content of education and applying it to a new linked dataset of four-year college graduates in a U.S. state.

The findings show that this form of alignment varies substantially and is consequential. Skill-task match varies substantially within majors, occupations, and common credential-to-occupation pathways, indicating that broad educational and occupational categories obscure meaningful heterogeneity in worker–job fit. The earnings return to skill utilization also extends beyond occupation-specific training: college-acquired skills are portable and rewarded across multiple labor-market destinations, and play an important role in explaining disparities among college graduates – especially between Black and White BA holders. The size of the earnings premium to skill-task match is notable precisely because it is not backed by the same linkages that organize conventional school-to-work linkages. Unlike tightly linked pathways, which may be reinforced by licensing regimes, professional categories, and recognizable hiring pipelines, transferable skill alignment is less credentialized and immediately legible to employers. Yet the results show that this less formalized form of alignment nevertheless carries substantial economic value.

There are several important limitations to these analyses. First, the data provide rare granularity into the relationship between schooling and employment, but are limited to one state's public college system and labor market. Patterns of match may differ in private colleges, in colleges with stronger liberal-arts orientations, and in the two-year sector, where vocational programs may map more directly onto common occupational pathways. Second, the estimated earnings returns exclude graduates whose employment is not observed in the linked data. In particular, workers who obtain out-of-state employment are not covered by in-state UI earnings records. To the extent that individuals migrate out of state to pursue better-matched employment, this could attenuate my estimated earnings returns to match, but it is difficult to estimate the extent of the bias. Relatedly, I am unable to examine match for employers that do not report occupations for their workers. These employers are disproportionately smaller firms (Appendix C), and common credential-to-occupation pathways may be especially limited in precisely these settings.

Third, I use occupations rather than jobs to infer work tasks and to determine skill-task match. Yet, such an approach could miscategorize individuals due to heterogeneity in task requirements across jobs within occupational categories (Martín-Caughey, 2021; Han and Cheng, 2026). A worker who appears mismatched to the overall task profile of an occupation may be well matched to a specialized position within that occupation, and vice versa. For example, a statistics major obtaining work as a biology researcher that requires expertise in biostatistics is highly skill-task aligned, even though my measure of skill-task match — at the level of the occupation — would mask this alignment. Future work linking transcripts to job postings, resumes, or employer-side task descriptions could distinguish within-job skill diversity from between-job heterogeneity inside occupations. Finally, my framework does not provide causal identification. While I control for a large set of college and pre-college covariates in my models, unobserved motivation, networks and family resources could explain the match premium that I document. Notwithstanding these limitations, this study provides the first granular view of the link between schooling and employment and of the transferability of higher education skills, and offers a first step toward understanding the earnings consequences of these dynamics.

8 Conclusion

Recent debates over higher education accountability increasingly ask whether programs deliver labor-market value. Accreditation reform has become a key site of this debate, with federal policy discussions placing renewed emphasis on whether institutions provide high-quality programs. These debates often presume that the labor-market value of higher education can be judged by whether programs prepare students for specific jobs. This paper suggests a more complicated picture. A large part of higher education's value may lie not in narrow occupational preparation, but in providing students with transferable skills that can be mobilized and rewarded across multiple job domains. Skill specificity is one component of the return to higher education; judging programs only by single-occupation placement risks undervaluing broad forms of undergraduate training.

At the same time, showing that skills are highly transferable across economic domains also reveals an important dimension of inequality. Graduates do not benefit equally from the skills they acquire in college. One advantage of tightly defined credential-to-occupation linkages is that large numbers of graduates with these credentials are able to match into skill-utilizing jobs. Even if broader skill alignment is an important component of earnings, broader skill utilization and transferability necessarily come with higher labor market frictions. Many graduates in lower-linked fields may simply not know what labor market destinations would constitute a good match. These frictions are compounded by resource disadvantages of lower-income groups, group differences in job search behavior, and hiring discrimination. Efforts to reduce postsecondary earnings gaps are therefore likely to be incomplete if they focus only on access to college, choice of major, or course-taking, without attention to the process through which graduates find skill-aligned work. Colleges, in particular, may have a role to play beyond the classroom: strengthening career services, alumni networks, and employer partnerships can help ensure that skills students acquire translate into better labor-market outcomes. At the same time, structural constraints—including family resources and bridging funds—limit what colleges can achieve without broader reform.

The framework developed here also opens broader questions about how skill investments matter over the life course. Over time, skill alignment is likely to be dynamic. Workers accumulate skills through jobs as well as through schooling (Gibbons and Waldman, 2004; Kalleberg and Mouw, 2018), and the relevance of those skills changes as technology, firms, and occupations evolve. Future research could therefore conceptualize match not only as the relationship between a graduate's transcript and job, but as the relationship between a worker's accumulated skill history and the changing task structure of available work.

Such a dynamic perspective could also recast skill-task match as a way to understand labor-market resilience. Job displacement may be less costly for workers whose skills align with multiple possible labor-market destinations, and more costly for workers whose skills are tied to a narrow set of tasks, firms, or occupational pathways. A task-based measure of match could therefore help identify not only who is well matched in their current job, but also who has viable alternative matches elsewhere in the labor market. In this sense, skill-task match provides a framework for understanding the capacity of workers to adapt as jobs, technologies, and task demands change.

Appendices

A Course-to-work-activity pipeline funnel

Here, I describe in more detail certain parts of the catalog processing that produces each student's skill profile. The processing proceeds in four stages: (1) catalog ingestion (raw course descriptions scraped from each college's online catalog); (2) administrative-record matching (catalog course → transcript course via exact code match, then fuzzy title match, then AI-assisted verification); (3) keyword filter (rule-based removal of administrative/logistical sentences); (4) LLM classification for skill-relevant sentences. Note that only course catalog text was passed through an LLM to tag for skill-relevant sentences in the course description; no student-level data was ever submitted to an LLM. I describe stages 3 and 4 in more detail below.

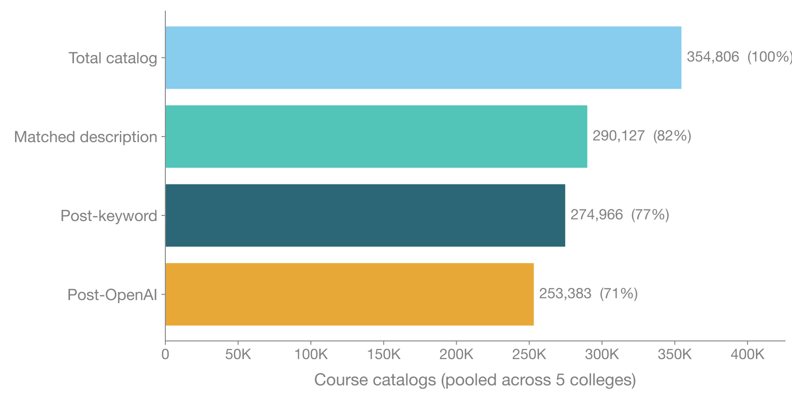

Figure A.1: Course-to-work-activity pipeline funnel: share of courses surviving each filter stage, pooled across all five colleges. Bars report cumulative retention from catalog ingestion through final work activity scoring.

A.1 Stage 3: keyword filter

The keyword filter drops any sentence containing one of the following terms (case-insensitive regex match, pipe-separated):

```
prerequisite|offered:|course details|myplan|instructors|offered|in
a sequence|open only to|no initial knowledge|credit|concurrent
enrollment|no auditors|required|meets professional
standards|mandatory|restricted|sequence|continuation
of|majoring|study abroad|for students|high school
```

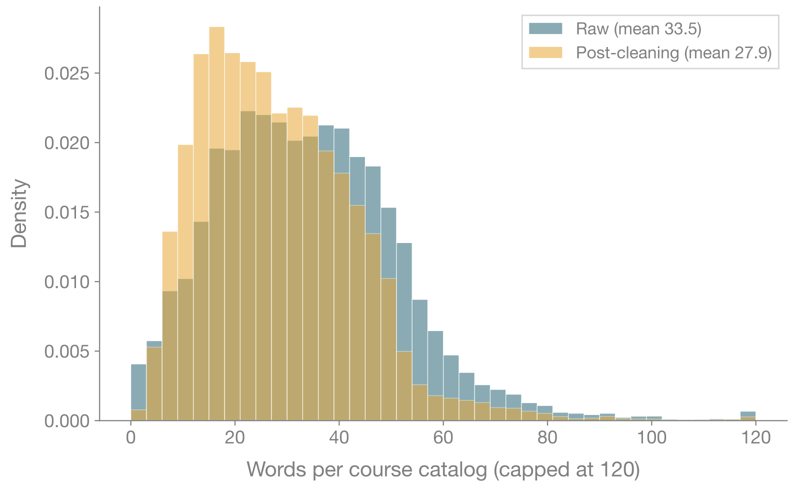

Figure A.2: Distribution of cleaned-paragraph word counts per course, pooled across all five colleges. The vertical line marks the 5-word minimum-length cutoff used to drop courses with insufficient surviving content.

The list was developed iteratively on a calibration sample to identify boilerplate sentences (prerequisites, scheduling, enrollment policies) that systematically inflate similarity to education- and pedagogy-related work activities.

A.2 Stage 4: LLM skill-relevance classifier

Surviving sentences are submitted to OpenAI's GPT-4o-mini for a binary skill-relevance classification. The system prompt is:

> You are an internal tool that classifies individual sentences extracted from university course descriptions (syllabi / catalog entries).
>
> For each input sentence, determine `skill_relevance`: does it describe skill-relevant content (course topics, methods, concepts, competencies) or logistics/prerequisites (enrollment info, credits, scheduling, prerequisites)?
>
> - 1 = skill-relevant course content
> - 0 = logistics, prerequisites, or administrative information
>
> *Rules:* Return EXACTLY one digit (`skill_relevance`); if ambiguous, lean toward 1 (keep it).

Example sentences classified as not skill-relevant. The following ten illustrative sentences received `skill_relevance = 0` and were dropped before embedding:

1. "Prerequisite: MATH 124 or equivalent placement."
2. "This course is offered every fall and spring quarter."
3. "Students must obtain instructor permission before registering."
4. "Final grade is based on three midterm exams (60%) and a final paper (40%)."
5. "Counts as a writing-intensive (W) credit."
6. "Class meets Tuesdays and Thursdays from 1:30 to 3:20 PM."
7. "Offered jointly with the Department of Economics."
8. "This course satisfies the quantitative reasoning (QSR) general education requirement."
9. "Restricted to students majoring in Computer Science with junior or senior standing."
10. "Course materials provided through Canvas; no required textbook."

B Additional sample descriptives

Table B.1: Distribution of the raw skill–task match score for BA completers. Columns report the sample size, mean, standard deviation, and key percentiles. The raw match score is the cosine similarity between a student's transcript-derived work-activity vector and their occupation's work-activity vector, prior to within-occupation standardization used in the regression analyses.

Sample	N	Mean	SD	p10	p25	p50	p75	p90
BA completers	32,694	0.355	0.266	0.031	0.125	0.300	0.585	0.757

Table B.2: Skill-Task Match by Common Match indicator.

	N	Mean	SD	p25	p50	p75
Common Match	8,438	0.609	0.228	0.485	0.671	0.779
Uncommon Match	24,256	0.267	0.216	0.092	0.218	0.400
Correlation (Common Match, Skill-Task Match): 0.563						

Table B.3: Sample descriptives by quartile of skill–task match. Columns compare students in the bottom and top quartiles of the (within-sample) Skill-Task Match distribution. Rows report demographic composition and pre-college academic preparation. Cell entries are percentages for categorical variables and means for continuous variables.

Characteristic	Quartile of Skill-Task Match	
	Bottom quartile	Top quartile
Female	56.5%	53.8%
Black	3.3%	1.4%
Hispanic	6.4%	5.1%
White	68.4%	70.7%
Asian	17.8%	18.2%
FRPL (ever)	20.6%	16.0%
Final GPA	3.50	3.60
Standardized test score	0.12	0.33

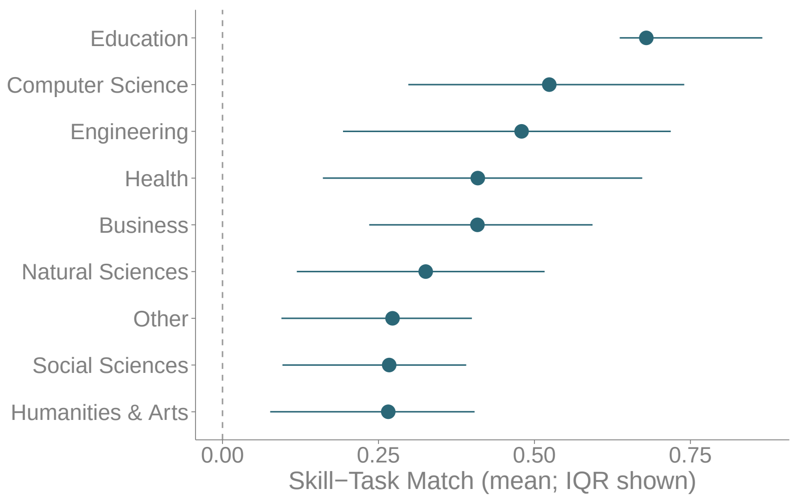

Figure B.1: Mean skill-task match score, by field of study (BA completers). Fields with weaker school-to-work linkages (liberal arts, humanities, social sciences) have the lowest mean match scores; fields with more structured pathways (health, engineering, business) have the highest. "Other" bundles low-headcount fields: architecture, protective services, family and consumer sciences, legal studies, other technical, communications and journalism, agriculture, public administration, and professional/culinary services.

C Sample balance

Table C.1 compares BA completers with positive UI-covered earnings to those with no observed earnings. The two columns are substantively similar on demographic and pre-college covariates; the No-earnings group has zero employer-reported SOC coverage by construction. Table C.2 similarly compares workers with and without an employer-reported SOC code among those with positive earnings.

Table C.1: Covariate balance for core sample with versus without UI earnings.

	Has earnings	No earnings
Female	56.4%	57.2%
Black	2.9%	3.0%
Hispanic	7.7%	5.4%
White	65.9%	68.5%
FRPL (ever)	23.2%	20.1%
Final HS GPA (mean, SD)	3.53 (0.43)	3.56 (0.44)
Standardized test score	0.15	0.32
Employer reported SOC code	71.0%	0.0%

Table C.2: Covariate balance for core sample with observed earnings, with and without SOC code.

	Has employer-reported SOC	No employer-reported SOC
Student characteristics		
Female	55.0%	59.9%
Black	2.5%	3.8%
Hispanic	6.1%	11.5%
White	68.6%	59.3%
FRPL (ever)	18.4%	34.8%
Final HS GPA (mean, SD)	3.54 (0.43)	3.52 (0.43)
Standardized test score	0.20	0.04
Employer characteristics		
Mean log earnings	11.31	11.23
Employer 10,000+ employees	26.7%	20.3%

D Robustness analyses

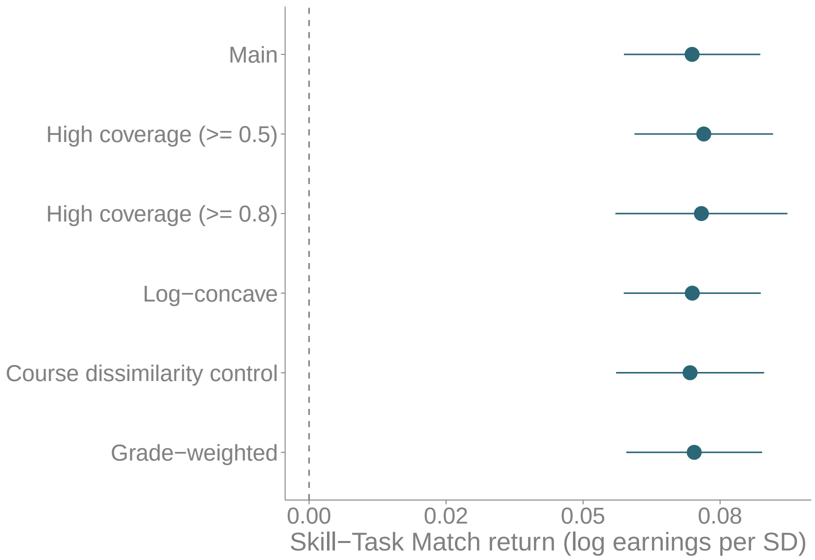

Figure D.1: Robustness of the main specification to alternative modeling choices. Restricting the sample to high-coverage students (those for whom at least 80% of transcript courses survive the keyword and LLM filters and are successfully embedded against the work-activity space) addresses the concern that systematic missingness in course-text content could be biasing the main results. Replacing the linear sum aggregation of work-activity scores with a concave $\log(1 + x)$ transformation models diminishing marginal returns to repeated exposure to similar course content—e.g., the marginal informational value of a fifth quantitative course is smaller than that of a second. In both robustness checks, the point estimates and 95% confidence intervals show that the earnings return to skill-task match is essentially unchanged.

E A major-level benchmark

The preceding argument treats transcript-derived skill-task match as a measure of the alignment between what graduates learned in college and the tasks they perform at work. Yet graduates enter the labor market not only with skills, but also with credentials. Declared majors organize course-taking and operate as socially recognized categories that help employers infer capacities, expertise, and appropriate job matches (Spence, 1973; Arrow, 1973; Bills, 2003; Rivera, 2011). As a result, graduates with high transcript-derived skill-task match may also hold jobs that align with the average skill profile associated with their major. If earnings rewards operate mainly through this broader category-level alignment, then the apparent earnings association with individual skill-task match may reflect major-implied alignment rather than the specific content of students' transcript trajectories.

I assess this possibility by comparing individual skill-task match to a parallel measure of major-task match: the alignment between the task content of the graduate's job and the average transcript-derived skill profile associated with the graduate's major. This benchmark is theoretically important because majors make some forms of worker–job fit more legible than others. Categories help audiences interpret actors and determine which criteria should be used to evaluate them (Zuckerman, 1999; Hannan et al., 2007; Leung, 2014). In labor markets, a job-seeker's major can perform this classificatory work by making some claims of fit more credible than others (Freidson, 1994; Weeden, 2002). From this perspective, the earnings association between individual skill-task match and earnings should be substantially reduced once major-task match is taken into account.[12]

At the same time, majors are useful but incomplete summaries of skill acquisition. The U.S. credential landscape is fragmented, and school-to-work linkages are comparatively weak relative to coordinated market economies (Hall and Soskice, 2001). Students within the same major often complete different course sequences, while students in different majors may acquire overlapping skills through electives, general-education requirements, minors, projects, and cross-field coursework (Han et al., 2023; Kizilcec et al., 2023). Course-level investments may also become visible beyond the formal credential title: graduates can highlight specific courses, tools, or portfolios in resumes and interviews (Kreisman et al., 2021), and employers may use skill assessments or early

[12]Empirically, Manuel (2024) measures the similarity between the skill requirements of an individual's occupation and the knowledge requirements of occupations closely associated with that individual's field of study. That paper finds an important payoff to major-occupation knowledge alignment, but cannot determine whether this alignment operates net of specific educational investments. The present analysis builds on this logic by asking whether the association between individual skill-task match and earnings is attenuated after accounting for major-task match, measured as the alignment between a graduate's job and the average transcript-derived skill profile associated with the graduate's major.

job performance to observe competencies that degrees and majors do not fully capture (Piopiunik et al., 2020; Sigelman et al., 2024). From both signaling and human-capital perspectives, then, transcript-derived skills may provide information about worker–job fit that broad credentials do not fully capture (Spence, 1973; Arrow, 1973; Stiglitz, 1975; Bills, 2003; Becker, 1962).

I therefore treat major-task match as a category-level benchmark for evaluating the interpretation of individual skill-task match. If the earnings association between individual skill-task match and earnings is substantially attenuated after accounting for major-task match, this would suggest that the observed association primarily reflects major-implied alignment. If the association remains substantial, this suggests that transcript-derived skill portfolios contain labor-market information beyond the implied skill profile of field of study.

To assess this possibility, I compare individual skill-task match to major-task match: the alignment between a graduate's job tasks and the average skill profile implied by their field of study. I measure major-task match using national course-skill profiles from Sabet et al. (2024), who infer major-level skill profiles from more than three million course syllabi at nearly three thousand U.S. higher-education institutions. For each graduate, I assign the national skill profile corresponding to their field of study and compute its cosine similarity to the task profile of their occupation. This score captures how closely the graduate's job aligns with the broader skill profile implied by their major, rather than with their own transcript trajectory. I then estimate earnings models that include individual skill-task match and major-task match separately and jointly.

Table E.1: Earnings returns to skill–task match versus a major-level skill profile constructed from national course-syllabus data (Sabet et al., 2024). Model 1 includes Average Major Match only, Model 2 includes Skill-Task Match only, and Model 3 includes both jointly. Standard errors are shown in parentheses. All specifications include pre-college controls, college-level controls, high-school fixed effects, and occupation fixed effects.

	Model 1	Model 2	Model 3
Skill-Task Match	–	0.070 (0.006)	0.071 (0.007)
Average Major Match	0.022 (0.004)	–	-0.002 (0.005)

Table E.1 reports the results. Major-task match is positively associated with earnings when entered in a model without skill-task match (conferring a .022 log increase in earnings per SD increase), suggesting that alignment between job tasks and the average skill profile associated with a graduate's field of study is rewarded in the labor market. However, once major-task match is entered in a model alongside skill-task alignment, the coefficient on individual skill-task match remains substantial, while the coefficient on major-task match attenuates to zero. This pattern indicates that the earnings return to skill-task alignment is not reducible to the skill profile implied by field of study. Instead, transcript-derived skill portfolios contain labor-market-relevant information beyond what can be inferred from a graduate's major.

I also construct a major-level measure of match internally from the course catalogs in my data: for each major, I average the work-activity scores across all courses associated with that major, then compute the cosine similarity between this major-level skill profile and the task profile of the graduate's occupation. These results, presented in Table E.2, are consistent with those in Table E.1.

Table E.2: Earnings returns to skill–task match versus a major-level skill profile constructed by averaging work-activity scores across all courses associated with each major in the catalog data. Model 1 includes Average Major Match only, Model 2 includes Skill-Task Match only, and Model 3 includes both jointly. Standard errors are shown in parentheses. All specifications include pre-college controls, college-level controls, high-school fixed effects, and occupation fixed effects.

	Model 1	Model 2	Model 3
Skill-Task Match	–	0.069 (0.006)	0.074 (0.008)
Average Major Match	0.036 (0.006)	–	-0.007 (0.007)

References

AACU (2010). Employers' views on college learning in the wake of the economic downturn. Technical report, Association of American Colleges and Universities, Washington, DC.

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

Altonji, J. G., Kahn, L. B., and Speer, J. D. (2014). Trends in earnings differentials across college majors and the changing task composition of jobs. *American Economic Review*, 104:387–393.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics*, 2(3):193–216.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323.

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

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70:9–49.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94:991–1013.

Bills, D. B. (2003). Credentials, signals, and screens: Explaining the relationship between schooling and job assignment. *Review of Educational Research*, 73:441–469.

Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. Wiley, New York.

Bleemer, Z. and Mehta, A. (2023). College major restrictions and student stratification. Working paper, Princeton University, Bendheim Center for Finance. URL: https://bcf.princeton.edu/wp-content/uploads/2023/01/Restrictions_Paper.pdf.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, pages 436–455.

Bol, T., Eller, C. C., Werfhorst, H. G. V. D., and DiPrete, T. A. (2019). School-to-work linkages, educational mismatches, and labor market outcomes. *American Sociological Review*, 84:275–307.

Bowen, W. G. and Bok, D. (1998). *The Shape of the River. Long-Term Consequences of Considering Race in College and University Admissions.* ERIC.

Brand, J. E. and Halaby, C. N. (2006). Regression and matching estimates of the effects of elite college attendance on educational and career achievement. *Social Science Research*, 35:749–770.

Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31:223–243.

Breen, R., Mood, C., and Jonsson, J. O. (2016). How much scope for a mobility paradox? the relationship between social and income mobility in Sweden. In *Income Inequality and Social Mobility: Sociological and Economic Perspectives*.

Béduwé, C. and Giret, J.-F. (2011). Mismatch of vocational graduates: What penalty on French labour market? *Journal of Vocational Behavior*, 78:68–79.

Cheng, S., Chauhan, B., and Chintala, S. (2019). The rise of programming and the stalled gender revolution. *Sociological Science*, 6:321–351.

Davies, S. and Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social Forces*, 75:1417–1438.

de Werfhorst, H. G. V. (2002). Fields of study, acquired skills and the wage benefit from a matching job. *Acta Sociologica*, 45:286–303.

Deming, D. and Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

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

DiPrete, T. A., Eller, C. C., Bol, T., and de Werfhorst, H. G. V. (2017). School-to-work linkages in the United States, Germany, and France. *American Journal of Sociology*, 122:1869–1938.

Forster, A. G. and Bol, T. (2018). Vocational education and employment over the life course using a new measure of occupational specificity. *Social Science Research*, 70:176–197.

Freidson, E. (1994). *Professionalism reborn: Theory, prophecy, and policy*. Polity Press, Cambridge, UK.

Friedman, S. and Laurison, D. (2019). The class ceiling: Why it pays to be privileged.

Gathmann, C. and Schönberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, 28(1):1–49.

Gerber, T. P. and Cheung, S. Y. (2008). Horizontal stratification in postsecondary education: Forms, explanations, and implications. *Annual Review of Sociology*, 34:299–318.

Gibbons, R. and Waldman, M. (2004). Task-specific human capital. *American Economic Review*, 94(2):203–207.

Goldin, C. and Katz, L. F. (2008). *The race between education and technology*. Harvard University Press, Cambridge, MA. Paperback edition 2010.

Grodsky, E. (2007). Compensatory sponsorship in higher education. *American Journal of Sociology*, 112:1662–1712.

Hall, P. A. and Soskice, D. (2001). *Varieties of capitalism: The institutional foundations of comparative advantage*. Oxford University Press, Oxford, UK.

Han, S. and Cheng, S. (2026). Skill diversification beyond high-paying jobs. *American Journal of Sociology*.

Han, S., LaViolette, J., Borkenhagen, C., McAllister, W., and Bearman, P. S. (2023). Interdisciplinary college curriculum and its labor market implications. *Proceedings of the National Academy of Sciences*, 120:e2221915120.

Hannan, M. T., Pólos, L., and Carroll, G. R. (2007). *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton University Press, Princeton, NJ.

Hauser, R. M. and Warren, J. R. (1997). Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology*, 27:177–298.

Hout, M. (1984). Status, autonomy, and training in occupational mobility. *American Journal of Sociology*, 89:1379–1409.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4):453–479.

Johnson, M. K., Sage, R. A., and Mortimer, J. T. (2012). Work values, early career difficulties, and the U.S. economic recession. *Social Psychology Quarterly*, 75:242–267.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990.

Kalev, A., Dobbin, F., and Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71(4):589–617.

Kalleberg, A. L. (2008). The mismatched worker: When people don't fit their jobs. *Academy of Management Perspectives*, 22:24–40.

Kalleberg, A. L. and Marsden, P. V. (2013). Changing work values in the United States, 1973–2006. *Social Science Research*, 42:255–270.

Kalleberg, A. L. and Mouw, T. (2018). Occupations, organizations, and intragenerational career mobility. *Annual Review of Sociology*, 44:283–303.

Kanter, R. M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, 82:965–990.

Kizilcec, R. F., Baker, R. B., Bruch, E., Cortes, K. E., Hamilton, L. T., Lang, D. N., Pardos, Z. A., Thompson, M. E., and Stevens, M. L. (2023). From pipelines to pathways in the study of academic progress. *Science*, 380:344–347.

Kmec, J. A. (2007). Ties that bind? Race and networks in job turnover. *Social Problems*, 54:483–503.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132:665–712.

Kreisman, D., Smith, J., and Arifin, B. (2021). Labor market signaling and the value of college: Evidence from resumes and the truth. EdWorkingPaper 21-423, Annenberg Institute, Brown University.

Laurison, D. and Friedman, S. (2016). The class pay gap in higher professional and managerial occupations. *American Sociological Review*, 81:668–695.

Laurison, D. and Friedman, S. (2024). The class ceiling in the united states: Class-origin pay penalties in higher professional and managerial occupations. *Social Forces*.

Leung, M. D. (2014). Dilettante or Renaissance person? How the order of job experiences affects hiring in an external labor market. *American Sociological Review*.

Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 4, pages 283–326. Elsevier.

Levels, M., der Velden, R. V., and Allen, J. (2014). Educational mismatches and skills: new empirical tests of old hypotheses. *Oxford Economic Papers*, 66:959–982.

Li, X. and Lu, Y. (2023). Education–occupation mismatch and nativity inequality among highly educated U.S. workers. *Demography*, 60(1):201–226.

Liu, Y. and Grusky, D. B. (2013). The payoff to skill in the third industrial revolution. *American Journal of Sociology*, 118(5):1330–1374.

Lovenheim, M. F. and Smith, J. (2023). Returns to different postsecondary investments: Institution type, academic programs, and credentials. In *Handbook of the Economics of Education*. Elsevier.

Lu, Y., Li, X., and Ciocca Eller, C. (2025). Underemployed and penalized: Education–occupation mismatch and racial/ethnic inequality among highly educated workers. *Sociology of Education*, 99(1):59–81. DOI: 10.1177/00380407251356275.

Manuel, N. (2024). Does the labour market value field of study specific knowledge? an alignment score based approach. *Economics of Education Review*, 101:102561.

Martín-Caughey, A. (2021). What's in an occupation? Investigating within-occupation variation and gender segregation using job titles and task descriptions. *American Sociological Review*.

McTague, T., Stainback, K., and Tomaskovic-Devey, D. (2009). An organizational approach to understanding sex and race segregation in U.S. workplaces. *Social Forces*, 87:1499–1527.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Columbia University Press.

Moss-Pech, C. (2025). *Major Trade-Offs: The Surprising Truths about College Majors and Entry-Level Jobs*. University of Chicago Press, Chicago.

Mountjoy, J. (2024). Marginal returns to public universities. Technical report.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, pages 693–709.

Pager, D., Bonikowski, B., and Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American Sociological Review*, 74:777–799.

Pager, D. and Pedulla, D. S. (2015). Race, self-selection, and the job search process. *American Journal of Sociology*, 120:1005–1054.

Piopiunik, M., Schwerdt, G., Simon, L., and Woessmann, L. (2020). Skills, signals, and employability: An experimental investigation. *European Economic Review*, 123:103374.

Poletaev, M. and Robinson, C. (2008). Human capital specificity: Evidence from the dictionary of occupational titles and displaced worker surveys, 1984–2000. *Journal of Labor Economics*, 26(3):387–420.

Rivera, L. A. (2011). Ivies, extracurriculars, and exclusion: Elite employers' use of educational credentials. *Research in Social Stratification and Mobility*, 29:71–90.

Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American Sociological Review*, 77:999–1022.

Robst, J. (2007). Education and job match: The relatedness of college major and work. *Economics of Education review*, 26:397–407.

Roksa, J. and Levey, T. (2010). What can you do with that degree? College major and occupational status of college graduates over time. *Social Forces*, 89(2):389–415.

Sabet, A. J., Bana, S. H., Yu, R., and Frank, M. R. (2024). Course-Skill Atlas: A national longitudinal dataset of skills taught in U.S. higher education curricula. *Scientific Data*, 11:1086.

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31(2):831–880.

Sigelman, M., Fuller, J., and Martin, A. (2024). Skills-based hiring: The long road from pronouncements to practice. White paper, Burning Glass Institute and Harvard Business School Project on Managing the Future of Work.

Skaggs, S. (2009). Legal-political pressures and African American access to managerial jobs. *American Sociological Review*, 74(2):225–244.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87:355–374.

Stiglitz, J. E. (1975). The theory of "screening," education, and the distribution of income. *The American Economic Review*, 65:283–300.

Tomaskovic-Devey, D., Stainback, K., Taylor, T., Zimmer, C., Robinson, C., and McTague, T. (2006). Documenting desegregation: Segregation in American workplaces by race, ethnicity, and sex, 1966–2003. *American Sociological Review*, 71(4):565–588.

Weeden, K. A. (2002). Why do some occupations pay more than others? Social closure and earnings inequality in the United States. *American Journal of Sociology*, 108:55–101.

Witte, J. C. and Kalleberg, A. L. (1995). Matching training and jobs: The fit between vocational education and employment in the German labour market. *European Sociological Review*, 11:293–317.

Witteveen, D. and Attewell, P. (2017a). The earnings payoff from attending a selective college. *Social Science Research*, 66:154–169.

Witteveen, D. and Attewell, P. (2017b). Family background and earnings inequality among college graduates. *Social Forces*, 95:1539–1576.

Wolbers, M. H. J. (2003). Job mismatches and their labour-market effects among school-leavers in Europe. *European Sociological Review*, 19:249–266.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30(1):1–53.

York, H., Song, X., and Xie, Y. (2025). Gradationalism revisited: Intergenerational occupational mobility along axes of occupational characteristics. *American Journal of Sociology*, 130(4).

Zhou, X. and Pan, G. (2023). Higher education and the black–white earnings gap. *American Sociological Review*, 88(1):154–188.

Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, 104(5):1398–1438.