Occupational Skill Mixing Under Breakthrough GPTs*

Elmer Zongyang Li†

August, 2026

Click here for the latest version

Abstract

Occupations increasingly demand skills that previously belonged to other jobs—a rise in *skill mixing* that I document in the near-universe of U.S. online job listings from the early 2010s up to 2024 and also in the O*NET task data. This trend is especially pronounced in lower-skill occupations, and is shaped by mixed skills that originate in high-skill jobs and are biased toward computer and IT. I trace the rise to breakthrough general-purpose technologies (GPTs), measured from patents that are highly novel and broadly applicable. In a shift-share design instrumented with California-origin innovation, a one-standard-deviation rise in exposure raises skill mixing by about a third of its secular rise. Wages rise by roughly 14 log points with no loss of employment, and both the mixing and wage gains are larger for less-educated workers. Routine-task automation, when entered alongside breakthrough-GPT exposure, results in an insignificant skill mixing response and lowers wages and employment, implying the two are distinct. A multidimensional matching model estimated using the reduced-form evidence attributes about two-thirds of the rise in mixing to a deepening of skill complementarity across occupations, and the rest to the the convex cost of operating mixed-skill jobs. The same deepening is also the main force behind the changes in wage gap.

Keywords: skill demand, technological changes, occupations, multi-dimensional skill
JEL Codes: J21, J23, J24, J31, E24

*I am extremely grateful to Philipp Kircher, Julieta Caunedo, Michael Lovenheim, and Mathieu Taschereau-Dumouchel for their guidance and continuous support. I also want to thank Justin Bloesch, Xiaoming Cai, Gustavo de Souza, Lukas Freund, Fatih Guvenen, Paul Jackson, Nir Jaimovich, Joseba Martinez, Sephorah Mangin, Simon Mongey, Kristoffer Nimark, Xincheng Qiu, Michelle Rendall, Pascual Restrepo, Elisa Rubbo, Evan Riehl, Guillaume Rocheteau, Similan Rujiwattanapong, Anna Salomons, Seth Sanders, Florian Sniekers, Sergio Ocampo Díaz, Michael Waldman, and Shengxing Zhang for their inspiring comments. Also thanks to numerous seminar participants. All errors are my own.

†Li: International Monetary Fund, 700 19th Street NW, HQ1 9-167, Washington DC, 20431. E-mail: elmer.zongyangli@gmail.com. Website: www.elmerli.net.

The nature of work in the United States has undergone profound changes in recent decades. A large literature shows that technological change has reduced demand for routine tasks while raising the importance of non-routine cognitive and social skills (e.g., Autor, Levy, and Murnane 2003; Acemoglu and Autor 2011; Deming 2017). Studies of industrial robots provide a stark example of technology's potential to displace workers (Acemoglu and Restrepo 2020). However, as occupational skill demands evolve, it remains unclear whether employers are favoring a specific set of skills or seeking a broader range of skills. This distinction matters. If employers redesign jobs to require mixtures of different skills, indicating "skill mixing," this would indicate stronger complementarity and could raise labor market returns.

This paper studies *skill mixing*–the tendency for employers to bundle together skill demands that were previously seen in different occupations. I measure it using both Lightcast (formerly "Burning Glass"), a near-universe of U.S. online job postings, and the Occupational Information Network (O*NET), the Department of Labor's survey of occupational requirements.[1] Two families of measures organize the analysis. Counts *detect* mixing: for each occupation, I count the skills that are new to it in 2024 but were already demanded elsewhere in early 2010s. The *skill mixing index* measures its *degree*: how close an occupation's skill bundle is to one in which every skill is demanded equally. Counts establish that mixing is happening and reveal which skills move; the index—which is unaffected by postings simply growing longer or the skill taxonomy finer—carries the aggregate trend and the main estimates.[2]

I find that the composition of occupational skills has churned significantly. Nearly half of the skills an occupation listed in 2024 were absent from it in early 2010s, and almost all of these—above 94 percent—are mixed skills, already present in *other* occupations in early 2010s rather than genuinely new to the economy.[3] This pattern is especially pronounced in low-skill occupations, such as Operators and Elementary occupations.[4] Two stylized facts

[1] Lightcast scrapes over 220,000 online sources between 2010 and 2024 and parses vacancy text into a taxonomy of more than 33,000 skills. While Lightcast postings are not a random sample of all vacancies, several studies document that they closely track administrative benchmarks of aggregate vacancy flows and cross-sectional occupational distributions (Lightcast 2024; Tsvetkova et al. 2024). Following Kim, Merritt, and Peri (2024), I reduce redundancy by merging roughly 2,000 skills that appear similar using cosine similarity of skill embeddings. O*NET is the successor to the Dictionary of Occupational Titles and a primary source in the task literature; I use only descriptors from incumbent surveys for longitudinal consistency.

[2] The two families are highly correlated, with a correlation ranging from 0.70 to 0.86 for within cell changes at the commuting zone $\times$ demographic cell level (Appendix Table A3)), but they serve distinct roles.

[3] Three other categories appear in this decomposition: "retained skills" are those present in the occupation in both early 2010s and 2024; "new skills" are those present in 2024 that did not exist in any occupation in early 2010s; and "obsolete skills" are those present in early 2010s but absent by 2024. Unlike mixed skills, these categories capture the persistence, emergence, or disappearance of skills rather than their diffusion across occupations.

[4] I use the International Standard Classification of Occupations (ISCO), the taxonomy provided in the Lightcast data, and show the patterns are robust to the Census occupation classification.

further characterize the nature of skill mixing. First, mixing follows a clear hierarchy. A force-directed network of skill flows shows that the mixed skills in lower-skill destination occupations predominantly originate in high-skill occupations (such as Professionals) rather than in other lower-skill occupations. Second, the content of mixing is tilted toward general-purpose content—computer and analytical skills, information technology, marketing; by contrast, domain-specific and routine skills are less mixed.[5]

The skill mixing index strengthens the same picture. From the early 2010s to 2024, the index rises 5.1 points, which represents a 9.5 percentile growth in the 2011 cell distribution.[6] Independently, the same index built from O*NET incumbent surveys rises about 10 percentiles in its 2005 distribution.[7] The rise in skill mixing is not an artifact of firm composition or standard openings and closings in posting text. I test this using three sample restrictions: holding constant employers or employer-by-occupation pairs across windows, and trimming postings to their core text. Every series continues to rise under these restrictions. Further decomposition shows that the rise is within occupations and is not driven by employment reallocation or by worker composition.

What drives the rise in skill mixing? I examine the impact of breakthrough general-purpose technologies (GPTs)—innovations that are highly novel and broadly applicable. I construct a measure of occupational exposure to breakthrough GPTs using the universe of U.S. patents. Following Kelly et al. (2021), a patent is a breakthrough if its text departs sharply from prior patents yet anticipates later ones. Breakthroughs whose influence spans many technological fields, captured by the Hall, Jaffe, and Trajtenberg (2001) generality index, are then classified as GPTs. I link these patents to occupations using high-dimensional vector embeddings of patent texts and O*NET technology descriptors.[8] This definition is restrictive. Among cited patents, approximately one tenth are breakthroughs and two percent are breakthrough GPTs, but the latter are cited across more than twice as many technological fields as other well-cited breakthroughs. While standard breakthroughs remain concentrated in high-skill roles, breakthrough GPTs extend into lower-skill clerical, service, and elementary occupations.

To estimate the impact of breakthrough GPT exposure, I use a Bartik (1991) shift-share

[5] Two classifications give the same answer: the Lightcast Open Skills taxonomy, which groups skills into career areas such as "Information Technology" and "Finance," and the routine/non-routine (RNR) categories of the task-based literature.

[6] Because the cosine scale has no natural cardinal meaning, I report the path in percentile ranks of the 2011 distribution, following Autor, Levy, and Murnane (2003) and Deming (2017). By comparison, social-skill task inputs grew by about 12 percentiles of the 1980 distribution over the 32 years from 1980 to 2012 (Deming, 2017).

[7] O*NET releases resurvey only part of the occupation list; I read 2005, 2009, 2013, and 2018 so that most occupations are updated at least twice, and I use only incumbent-survey descriptors for longitudinal consistency.

[8] Patents are from the U.S. Patent and Trademark Office's PatentsView database. Section II details the construction; Appendix A.6 rebuilds exposure under alternative cosine cutoffs and discusses the results.

design for the years 2010 to 2024. Commuting-zone by demographic cells that started 2010 with different occupation mixes experience the same national breakthrough-GPT shocks, but with varying intensity.[9] I estimate the effects using 2SLS, instrumenting exposure with a California-recentered shift-share based only on California-assignee patents and year-demeaned across occupations within each year, following Borusyak and Hull (2023).[10] The exclusion restriction is that California-frontier innovation is uncorrelated with non-California labor-demand shocks, given these controls. This approach follows the standard leave-out or foreign-frontier logic used to separate technology and trade shocks in local labor markets.

The findings indicate that breakthrough GPTs raise skill mixing and wages without displacing workers. A one standard deviation increase in GPT exposure raises the skill mixing index by 1.89 points, representing about a third of the index's 5.1-point secular rise.[11] The detection measures move with it. The share of postings listing any mixed skill rises by 3.2 percentage points, and the log count among postings that already list one rises by 37.1 log points. The same one-standard-deviation exposure raises wages by 14.0 log points. These gains are most pronounced among less-educated workers, as the wage response drops from 22.0 log points for workers without a high school diploma to 4.9 log points for postgraduates, which represents a ratio of about four and a half to one. The mixing response is also largest among workers without a high school diploma. The ACS employment rate does not change, suggesting the effects run through the content of jobs, not the number of jobs. These effects of breakthrough GPTs are not an artifact of any single market or occupation. After omitting the five largest patent-producing states (CA, NY, TX, WA, MI), and after carrying out a leave-one-ISCO-out re-estimation as well as using alternative measures of GPTs and recentering, the findings for both skill mixing and wages remain consistent.

Racing breakthrough-GPT exposure against the leading alternative account—routine-biased technical change—shows that the two are not the same technology under different labels. I enter the Autor and Dorn (2013) routine-task intensity alongside the instrumented breakthrough-GPT stock.[12] Adding the competitor leaves the breakthrough-GPT estimates

[9]The unit is a commuting zone by sex, education, and experience cell by year over 2010 to 2024. The shares are fixed at 2010 American Community Survey (ACS) employment, so within-cell variation comes only from the national occupation shocks (Bartik 1991; Goldsmith-Pinkham, Sorkin, and Swift 2020; Borusyak, Hull, and Jaravel 2022; Adão, Kolesár, and Morales 2019).

[10]The estimating equation includes commuting-zone fixed effects, demographic-cell by year and state by year fixed effects, and a commuting-zone covariate vector. Standard errors are two-way clustered by commuting zone and state. Year-demeaning ensures unconditional exchangeability of occupation shocks within each year and removes the part of the California shock that is common to all occupations.

[11]The first stage is strong at the shock level. Clustering on the occupation shocks, the effective F ranges from 58 to 76 in the year-demeaned columns. Anderson–Rubin 95 percent intervals exclude zero on every posting-based mixing margin and on wages, while the employment interval brackets zero.

[12]The race retains the baseline fixed effects, the commuting-zone covariates, 2010 employment weights, and

where they were. The routine-task coefficients indicate displacement effects, contrasting with the impact of breakthrough GPTs. A one-standard-deviation increase reduces wages by approximately 1.7 log points and lowers the probability of any skill mixing by 0.4 percentage points. Furthermore, their loading on the mixing index is negative and imprecise.[13] Neither exposure affects the employment rate, but the two diverge on the employment level—routine-task exposure lowers it, while breakthrough GPTs raise it.

Prima facie, these responses are consistent with three explanations. Skills may have become stronger complements in production. The cost of designing and operating broad jobs may have changed. Worker skill supply may also have shifted. I develop a multidimensional matching model in which firms choose skill requirements by trading production gains from combining skills against a convex operating cost, in the spirit of Acemoglu (1999).[14] I calibrate the production parameters to the projected mixing responses, the occupational wage gap response, and labor market anchors.

The estimates point to a deepening of skill complementarity concentrated in the low-wage occupations. In elasticity terms, substitution across skills falls from 9.77 to 5.39 there, against 4.01 to 3.80 in the high-wage group. At the same time, the cost of operating mixed-skill designs becomes more convex, so that concentrating a job on a few skills is more expensive at the margin. That deepening accounts for about two-thirds of the rise in mixing the Section III estimates identify; rising operating costs account for most of the remainder. It is also the force behind the wage-gap compression. On its own it would compress the gap by roughly three-quarters again as much, and the cost channel gives back nearly half. The calibrated shift in worker skill supply has minimal impact throughout. In short, breakthrough GPTs raise the return to combining skills rather than specializing in one. That is why low-wage occupations mix most, and why their wages catch up.

Related literature. I study labor market dynamics focusing on the phenomenon of *skill mixing*. This work connects to the vast literature on long-term skill demand and technological change (e.g., Autor, Levy, and Murnane 2003; Goldin and Katz 2010; Acemoglu and Autor 2011; Acemoglu and Restrepo 2020). However, unlike studies that focus on demand that is either skill-biased or task-biased, I study skills in *mixtures* and document a structural shift where employers increasingly mix skills that were previously specific to other occupations.

two-way clustering by commuting zone and state. Once those are absorbed, the residualised exposures are negatively correlated at -0.05.

[13]Equality of the per-SD effects is rejected at $p < 0.001$ on every posting and wage margin.

[14]The departure from existing multidimensional matching models (Lindenlaub 2017; Lise and Postel-Vinay 2020; Ocampo 2022) is endogenous job design. Firms choose the skill intensity of each vacancy rather than taking an occupation profile as given.

These findings on the importance of skill recombination within occupations align with recent evidence that within-occupation variation in task content is a key margin of changing skill demand (Atalay et al. 2020; Freeman, Ganguli, and Handel 2020).[15]

My analysis contributes to the emerging literature that uses real-time posting data to study skill demand. Previous work shows that skill requirements are heterogeneous and predict firm performance (Atalay and Sarada 2020; Deming and Kahn 2018), while also rising during recessions (Hershbein and Kahn 2018). Other research uses these data to show that displaced workers face earnings losses when their skills lag (Braxton and Taska 2023) and to track the diffusion of "new work" and specific technologies across occupations (Kim, Merritt, and Peri 2024; Kalyani et al. 2025). The skill mixing is different: high-skill, non-routine content moves *down* the occupational ladder and raises pay there, without displacing the workers it reaches.

This paper also contributes to the literature on general-purpose technologies. Early work was largely theoretical (Helpman and Trajtenberg 1994); later evidence shows that GPT diffusion—for example the steam engine—is often slow, constrained by downstream adaptation (Crafts 2004). Recent research stresses that GPTs raise productivity not only directly but by inducing costly organizational change and complementary intangible investments (Brynjolfsson, Rock, and Syverson 2021). In ICT, that mechanism has been used to explain both the U.S. productivity advantage (Bloom, Sadun, and Reenen 2012) and shifts in labor demand (Katz and Murphy 1992; Autor, Levy, and Murnane 2003). Fewer studies measure how job requirements themselves reorganize as GPTs diffuse. I provide direct evidence that breakthrough GPTs mix skills from high-skill occupations into lower-skill jobs, and evaluate the consequences for wages and employment.

Theoretically, I interpret these findings using a multi-dimensional matching model with endogenous occupation design, extending the literature on matching with multidimensional heterogeneity (e.g., Lindenlaub 2017; Lise and Postel-Vinay 2020; Ocampo 2022). Recent literature on multidimensional matching (e.g., Yamaguchi 2012; Lindenlaub 2017; Lise and Postel-Vinay 2020) explores the assortativeness of matching and worker skill evolution. I extend this by examining firms' endogenous skill demand tradeoffs in a general equilibrium framework. By allowing firms to optimally design skill bundles subject to convex complexity costs following the spirit of Acemoglu (1999), the model rationalizes why breakthrough GPTs increase skill mixing.

The remainder of the article proceeds as follows. Section I documents the rise of skill mixing. Section II shows how I identify breakthrough GPTs in the patent record, and Section III

[15] Extracting task information from job ads, Atalay et al. (2020) reveal that major job content changes from 1950 to 2000 occurred within occupations, a trend that Freeman, Ganguli, and Handel (2020) find continues post-2000.

estimates their effects on skill mixing, wages, and employment. Section IV interprets the evidence through a multi-dimensional matching model, Section V quantifies the channels behind the rise in mixing, and Section VI concludes.

I The Mixing of Skills in the U.S. Labor Market

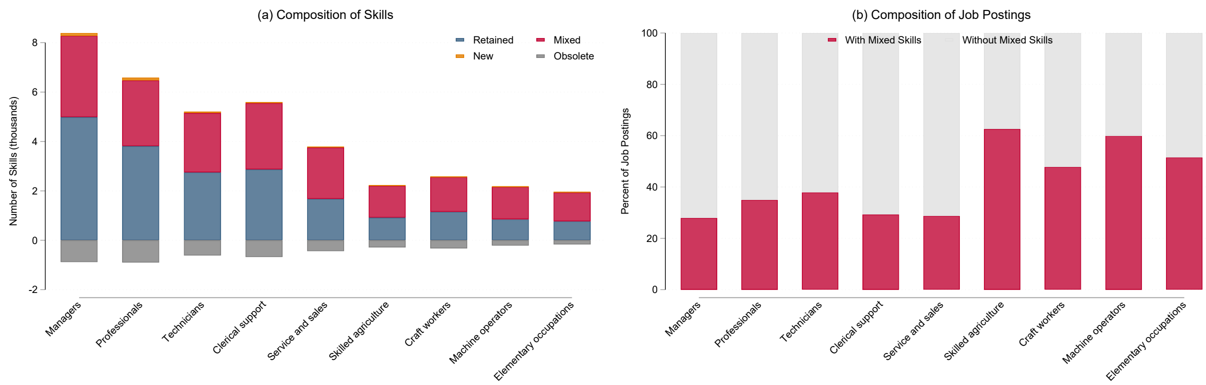
In this section I look at one aspect of the U.S. labor market from the early 2010s to 2024, namely whether and how skills cross occupation boundaries. I use vacancy data from Lightcast, together with ONET task data, and proceed in four steps. First, I determine what proportion of the skill requirements in an occupation in 2024 had not been there in the early 2010s to find out whether they are entirely new or originating from other occupations. Second, I identify which occupations act as sources of skill mixing and which serve as destinations, as well as examine the specific skill content. The third step involves measuring the degree of skill mixing using Lightcast and O*NET data to assess the proportionate use of different skills rather than just counting the number of mixed skills. Finally, I examine these measures year by year and decompose the trends. Robustness checks on the posting sample and measurement validity are presented throughout.

I.A Detecting Skill Mixing

Data. To detect and measure skill mixing, I primarily rely on the Lightcast database (formerly "Burning Glass Technologies"), spanning the period from 2010 to 2024. This dataset covers a near-universe of online job postings in the United States, compiled by systematically scraping over 220,000 online sources, including major job boards and corporate career pages (Lightcast 2024). Lightcast parses the raw text of millions of vacancies into detailed, high-frequency fields on employer demand (e.g., job title, location, occupation). Further, Lightcast applies a two-step deduplication process that identifies newly scraped ads and then compares fields within a 60-day window to remove repeated listings. While Lightcast does not provide a random sample of vacancies, benchmarking exercises indicate that its U.S. series and cross-sectional distributions align with official sources.[16]

[16]Using active postings, Lightcast reports that since 2013 it captures, on average, 92.6 percent of Job Openings and Labor Turnover Survey (JOLTS) job openings, and that the Lightcast and JOLTS series have a correlation of 0.87. It also reports strong alignment with Occupational Employment and Wage Statistics (OEWS) in the distribution of postings across SOC major groups (correlation 0.74) and across states (correlation 0.98) from May 2021 to May 2022. Lightcast notes that jobs not posted online are often in small businesses and union hiring halls, and the OECD documents that online postings track official distributions well but can remain systematically over- or under-represented in specific sectors, occupations, and regions (Lightcast 2024; Tsvetkova et al. 2024).

Figure 1: Composition of Skills in US Job Postings, 2010 to 2024



Notes: This figure plots the composition of skills in U.S. Lightcast postings from 2010 to 2024. Panel (a) splits skills in each major occupation group into four categories based on their emergence and persistence: retained, mixed, new, and declining. Mixed skills are those an occupation lists in 2024 that were absent from it in 2010 but present in other occupations. Calculations are at the four-digit ISCO level and then averaged to one-digit groups. Values are in thousands of skills, and declining skills sit below the horizontal axis. Panel (b) plots the share of postings in each one-digit ISCO group that list at least one mixed skill.

Lightcast also parses detailed skill requirements (tags) from raw vacancy text, and prior work uses this information to study trends in job skill demand (Deming and Kahn 2018; Hershbein and Kahn 2018; Braxton and Taska 2023). The database includes a taxonomy of over 33,000 continuously updated skills. To reduce redundancy, I consolidate about 2,000 near-duplicate skills using cosine similarity from embeddings, choosing the threshold based on the lowest similarity score within O*NET modules (Kim, Merritt, and Peri 2024). Two margins are used downstream: the count of mixed skill tags detects whether skills cross occupations, and the skill mixing index—built from four task-based intensities on the same tags—measures how evenly demand is spread.

Broad Pattern. How have occupational skill demands changed over time? To answer this question, I decompose employer skill requirements between 2010 and 2024 using Lightcast job posting data. The decomposition classifies skills at the 4-digit ISCO level and averages within 1-digit occupational groups (Figure 1). I group skills into four categories: (i) retained skills, which remain within the same occupation throughout this period; (ii) new skills, which emerge by 2024 but were not present in 2010; (iii) obsolete skills, which appear in 2010 but disappear by 2024; and finally, (iv) mixed skills, which appear in an occupation in 2024 but were not listed for that occupation in 2010 and instead appeared in other occupations in 2010.

The main finding of Figure 1 is that a large part of employer skill demand changes between 2010 and 2024 occurred through the mixing of existing skills across occupations. On average, 48.0 percent of all skills observed in occupations in 2024 were not present in the same occupation in 2010 (the union of "new" and "mixed" shares), and within this newly observed set, mixed skills (skills that appeared in other occupations in 2010) account for 95–99 percent — so mixed skills make up roughly 46.1 percent of all skills observed in 2024. While high-skill occupations demand the largest number of mixed skills, the share of mixing is higher in low-skill occupations.[17] Panel (b) reports the share of postings requiring mixed skills. Mixing is widespread and disproportionately affects lower-skill jobs, reaching 48 to 62 percent in elementary, operator, operator, and agriculture roles compared to 28 to 38 percent in high-skill occupations.

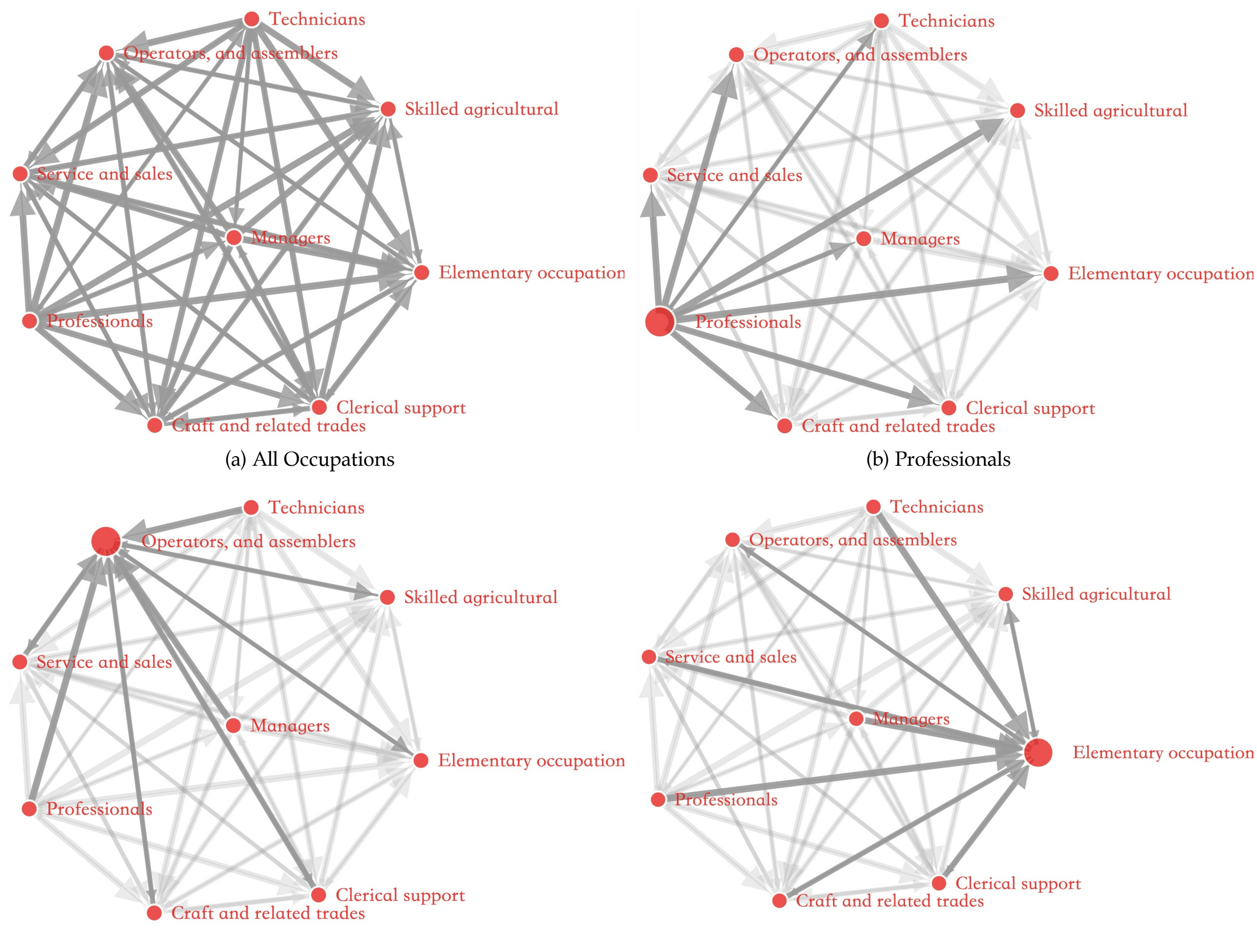
Robustness checks. The baseline results establish that skill mixing is the dominant entry channel for newly observed occupational skills. One may worry that this extensive count is an artifact of how skill presence is coded. Online Appendix Table A5 reconstructs the four-category decomposition using three sets of alternative definitions. First, sparse detections and increasing posting volumes could affect the mixing measure. If the threshold is raised to ten postings, with a requirement that a skill appears in at least 0.1 percent of an occupation's postings—a rule that does not change with volume growth—or if the sample is limited to occupations above the median 2010 posting volume, the overall mixed share of 2024 skills shifts to between 41.4 and 51.9 percent. Within the set of newly observed skills, mixed skills continue to account for 95 to 98.5 percent. Second, the choice of base year could artificially inflate later arrivals. Replacing the 2010 anchor with 2011 or 2012 lowers the overall mixed share to roughly 40 percent because more skills are already on record, but the mixed share of newly observed skills remains stable near 95 percent. Third, employer entry over the panel could mechanically generate skill mixing. Restricting the sample to employers present in both the early and late windows, or to employer–occupation pairs observed in both, leaves the mixed share between 42.7 and 46.7 percent. Across all alternative definitions, the overall mixed share of 2024 skills consistently remains at around 40 and 50 percent.

I.B Network and Content of Skill Mixing

To understand the skill mixing in the United States from the early 2010s to 2024, I conduct two analyses. First, I use a network analysis to examine which occupations serve as the main

[17] The granularity of the occupational classification affects these shares, as broader occupation groups would yield lower measured levels of mixing. However, online Appendix A.3 shows that the same patterns persist using an alternative 3-digit OCC classification. Across occupations, mixed skills remain substantial, averaging 49 percent for low-skill and 37 percent for high-skill occupations.

Figure 2: Network of Occupational Skill Mixing, 2010-2024



(a) All Occupations

(b) Professionals

(c) Operators and Assemblers

(d) Elementary Occupations

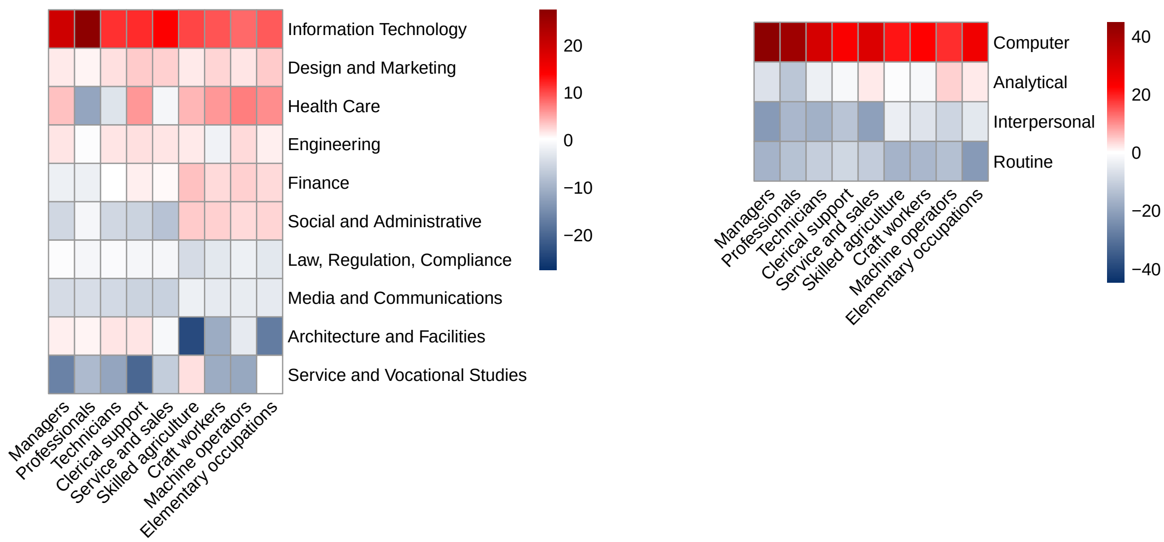
Notes: This figure plots occupational skill mixing in U.S. Lightcast postings from 2010 to 2024. Panel (a) is a force-directed network at the one-digit ISCO level. Each node is a major occupation group. Edge thickness is the count of skills that appear in one group in 2010 and in another in 2024. Panels (b)–(d) report the number of mixed skills that three destination groups absorb from other one-digit groups: Professionals (b), Operators and assemblers (c), and Elementary occupations (d).

sources of skill mixing and which act as destinations. Second, I examine the content of skill mixing. by tracking the composition of mixed skills within occupations across skill categories.

Network of skill mixing. The flow of skill mixing across occupational groups (Figure 2) panel (a) first visualizes the flow of skill mixing across occupational groups from the early 2010s to 2024 using a force-directed network.[18] Each node represents an ISCO one-digit occupation, and each directed edge captures the extent to which skills appearing in one

[18]The network layout is generated using a force-directed algorithm, which treats nodes as repelling particles and edges as springs that pull connected nodes together. The system simulates these forces to reach a stable equilibrium, so that occupations sharing more mixed skills are positioned closer, while those with fewer shared skills are pushed apart, which naturally reveals clustering patterns.

Figure 3: Relative Skill Mixing Intensity Across Categories



Notes: This figure plots relative mixing intensity across skill categories and occupations. Relative intensity is a category's share among mixed skills minus its share in the economy-wide skill stock. Only differences larger than one percentage point are shown. Panel (a) uses the Lightcast Open Skills taxonomy. Panel (b) maps skills to four Routine and Non-Routine groups (Analytical, Interpersonal, Computer, and Routine) under the three-layer assignment of Appendix A.2. Lightcast software skills map to Computer. The keyword lexicon of Hershbein and Kahn (2018) and Braxton and Taska (2023) covers the remaining named categories. Remaining skills are assigned to the nearest task-category centroid in embedding space only above a calibrated similarity threshold. Skills that clear no layer are excluded, so both shares in panel (b) are computed over RNR-classified skills.

occupation in 2010 also appear in others in 2024. The thickness of the edge corresponds to the count of skills. The network shows that high-skill occupations such as Professionals and Technicians are key sources of skill mixing. In contrast, low-skill occupations such as Elementary occupations, Skilled agriculture, and Craft and related trades appear more as recipients than sources.

Panel (b) focuses on Professionals, which emerge as a dominant "origin" of skill mixing. For example, over 4,600 skills originating from Professionals appear in Elementary occupations and among Operators and assemblers. However, there is significantly less mixing with high-skill peer occupations such as Technicians and Managers, with fewer than 2,600 skills. Panels (c) and (d) show that lower-skill occupations primarily function as "destinations." Operators and assemblers (panel c) absorb a large number of skills from Technicians and Professionals, suggesting that traditional industrial jobs are becoming more complex by increasingly requiring skills previously found in higher-skill roles. Similarly, Elementary occupations (panel d) act mainly as a destination, absorbing skills primarily from higher-skill roles and, to a lesser extent, from Clerical support.

Content of Skill Mixing. I next analyze the content of skill mixing by examining the specific categories of skills being mixed. The relative mixing intensity of different skill categories—the difference between a category's share within mixed skills and its share in the economy-wide skill stock—reveals a clear technological bias (Figure 3).[19] A positive value indicates that a skill type is disproportionately overrepresented in the mixed skills relative to its overall prevalence. Panel (a) categorizes skills based on the Lightcast Open Skills taxonomy.[20] Information Technology and Design and Marketing skills emerge as the two most highly mixed categories, exhibiting large positive values across different occupations. Low-skill occupations also integrate certain domain-specific skills, such as Health Care, Finance, and Social and Administrative competencies. In contrast, Architecture and Facilities, and Service and Vocational Studies display negative relative shares across all occupations.

To connect with the task-based literature, the right panel maps these patterns into four Routine/Non-Routine (RNR) groups following Acemoglu and Autor (2011): Analytical, Interpersonal, and Routine.[21] I additionally add a Computer group following Braxton and Taska (2023). I apply a three-layer rule to map Lightcast skill names into these groups (Appendix A.2). First, I assign remaining skills using the keyword lexicon from Hershbein and Kahn (2018) and Braxton and Taska (2023) (Appendix Table A2). Second, I assign any remaining skill to the nearest task-category centroid in the embedding space. I only make this assignment if the cosine similarity exceeds a 0.45 threshold, which delivers approximately 0.90 out-of-sample agreement with the first two layers.[22] Lastly, I report skills that fail all three layers as non-RNR domain skills. That residual is large because the rule leaves ambiguous tags out rather than forcing them into a category. Uncategorized tags account for 50.2 percent of the mixed-skill mass. Therefore, the RNR grouping categorize skills across the four task categories rather than across everything an employer writes down.

The RNR grouping confirms the technological bias of panel (a). Computer skills are the overwhelming driver of mixing across all occupations. Their share among mixed skills are 37 percentage points above their share of the economy-wide stock, and they are the

[19]Normalizing by the economy-wide skill stock filters out pure size effects across skill categories, ensuring that the measure captures skills that are disproportionately represented in mixed skills. The figure displays only those skills where the difference in relative share exceeds one percentage point.

[20]See the Lightcast Open Skills taxonomy documentation (https://lightcast.io/open-skills/categories). In this taxonomy, categories are broad groupings that map roughly to career areas (e.g., Information Technology, Finance, and Health Care).

[21]To maintain a feasible skill dimension, I separate the two non-routine skills (analytical and interpersonal) and consolidate the two routine skills (routine cognitive and manual) into a single category.

[22]I manually review the 100 most widely mixed skills that clear this embedding threshold. I keep the machine label if the plain meaning fits and withhold it otherwise. I leave withheld skills uncategorized rather than reassigning them by hand.

only group over-represented among skills that are mixed. Analytical skills mix about in proportion to its size, tilting positive only in lower-skill service and elementary occupations. In contrast, Routine skills fall 13 percentage points short of their stock share in every occupation group, and interpersonal skills are likewise under-represented by 17 percentage points. Taking stock, what diffuses into lower-skill occupations is mainly Computer and, within Lightcast's taxonomy, Information Technology and Design/Marketing—general-purpose skills, not routine ones.

I.C Measuring the Degree of Skill Mixing

I.C.1 The Skill Mixing Index

The count of mixed skills and the share of postings that carry one establish that skills are crossing occupational lines, and the network and content tell which skills move and where. What remains is to evaluate the *degree* of skill mixing. To do so, one can analyze the angular difference between the occupation's skill vector and the unit vector, on which different skill requirements are equivalent. The emphasis of skill mixing is on the proportionate use of different skills rather than the overall skill intensity. Building on this concept, I employ the cosine similarity between an occupation's skill vector and a multi-dimensional norm vector.[23]

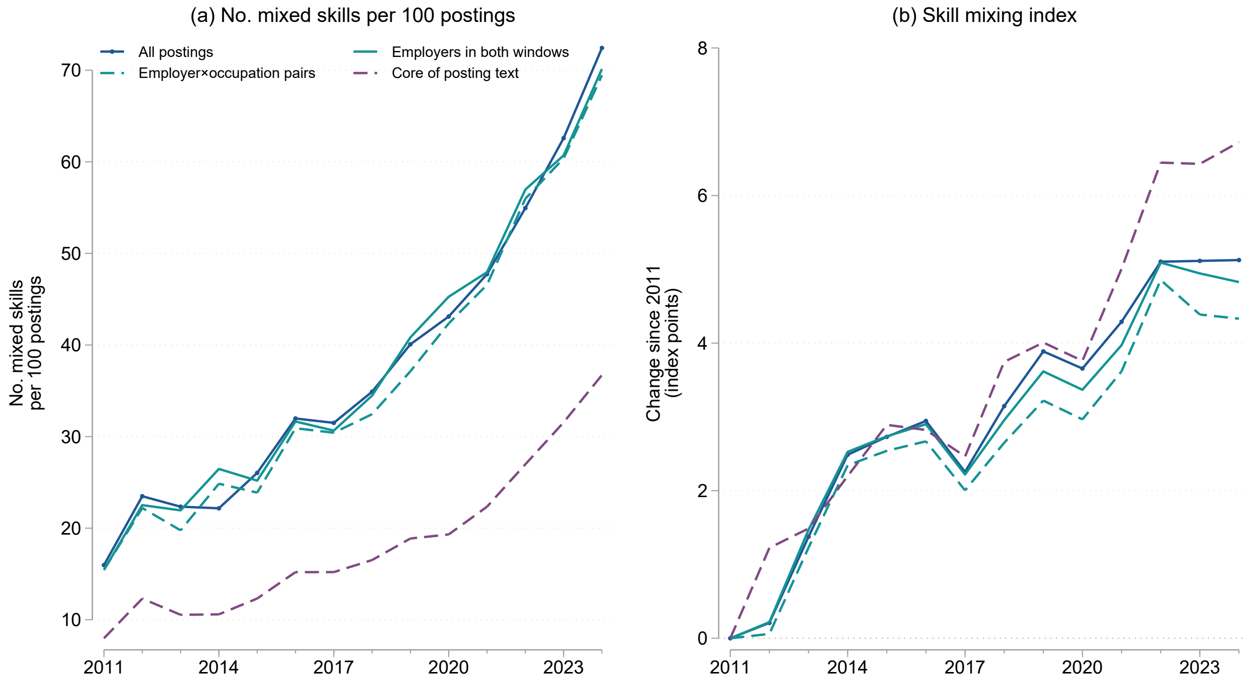
I give each occupation a vector of its four RNR skill intensities built as above, and normalize them by principal components and rescale linearly to the unit interval $[0, 1]$ so the four dimensions are cardinally comparable. Specifically, the skill mixing index for an occupation j in a K-dimensional space characterized by the skill intensity vector $\mathbf{y}^j$ is the cosine similarity between its skill vector and the norm $\hat{\mathbf{v}}$ in the skill space:

$$Mix(\mathbf{y}^j) = \frac{\mathbf{y}^j \cdot \hat{\mathbf{v}}}{||\mathbf{y}^j||, ||\hat{\mathbf{v}}||}, \quad \text{where } \hat{\mathbf{v}} = [1, 1, ..., 1]' \in \mathbb{R}^{K+}. \tag{1}$$

The mixing index as in equation (1) evaluates the multi-dimensional angular similarity between a skill vector $\mathbf{y}^j$ and the multi-dimensional norm $\hat{\mathbf{v}}$. As different skills in $\mathbf{y}^j$ get closer to each other, the value of $Mix(\mathbf{y}^j)$ will rise accordingly. There are three key advantages of the skill mixing index defined using cosine similarity that are worth noting. First, it easily accommodates occupations represented by multi-dimensional skills. Second, this index is independent of the length of the skill vector, and focuses on the proximity between a skill vector and the norm, which indicates the degree of skill mixing. Lastly, this measure is

[23] An angle-based measure is by no means the only measure of skill mixing, though it has the clearest graphical illustration of the trade-off among skills.

Figure 4: Trend in Skill Mixing, 2011–2024, under Alternative Posting Samples



Notes: This figure plots two measures of skill mixing in U.S. Lightcast postings from 2011 to 2024. Panel (a) is mixed skill tags per 100 postings. A mixed tag is one an occupation lists in 2024 that was absent from it in 2010 but present in other occupations. The navy line is the employment-weighted cell mean (16 in 2011, 72 in 2024). Panel (b) is the four-skill RNR mixing index of equation (1) at commuting-zone × four-digit occupation cells, employment-weighted, as the change since 2011. The navy line assigns tags by name (software titles, keyword lexicons, then an audited embedding layer) and leaves unnamed tags out. Both panels redraw three restricted samples: employers in both 2010–12 and 2022–24, employer-by-occupation pairs in both windows, and the core of the posting text after dropping the first and last fifty words. Weights are ACS hours-based employment, with 2024 using the 2023 ACS. Undrawn variants of the headline index are a fifth category for unassigned tags (3.2), the unweighted cell mean (5.6), posting-count weights (4.3), and the legacy forced-map index (6.3). 2010 is omitted because tagging is sparse.

inherently normalized, as the cosine of an angle in the first quadrant lies in $[0, 1]$. Reported results multiply the index by 100, and with four nonnegative skills it runs from 50 (a fully specialized bundle) to 100 (a fully balanced one).

The count of mixed skills and the skill mixing index line up closely. They correlate at 0.86–0.92 in levels and 0.70–0.86 in within-cell changes at the commuting-zone $\times$ demographic-cell level (Appendix Table A3). The rest of the paper measures degree with the index and uses the count as a detection margin.

I.C.2 The Trend, Early 2010s–2024

The analysis so far has compared two cross sections. This subsection tracks the year-by-year path of skill mixing from the early 2010s to 2024. Figure 4 reports two measures. The first measure is the average number of distinct mixed skill tags per posting, which detects mixing. The second measure is the skill mixing index, which measures its degree. I construct both series at the commuting-zone $\times$ four-digit occupation-cell level. Panel (a) plots the employment-weighted cell mean of mixed tags per 100 postings for each series. Panel (b) plots the employment-weighted cell-mean index as the change since 2011 on the 50–100 cosine scale.[24]

The trend. Figure 4 shows a substantial rise in skill mixing from the early 2010s to 2024. In Panel (a), postings carried 16 mixed skill tags per 100 in 2011 and 72 in 2024. This represents roughly a four-and-a-half-fold increase, and no single year accounts for the change. Panel (b) turns to the skill mixing index, which rises 5.1 points by 2024. The cosine scale has a limitation because one cannot comfortably treat it as cardinal. To address this limitation, Appendix Figure A2 maps the same path into percentile ranks of the 2011 cell distribution (see Autor, Levy, and Murnane (2003) and Deming (2017) for other examples). On that scale, the average cell is about 9.5 percentiles higher than its 2011 level.[25]. Despite the different units, the same qualitative pattern holds. Both margins record a rise in skill mixing.

I.D Measurement Validity and Decomposition

I.D.1 Measurement Validity

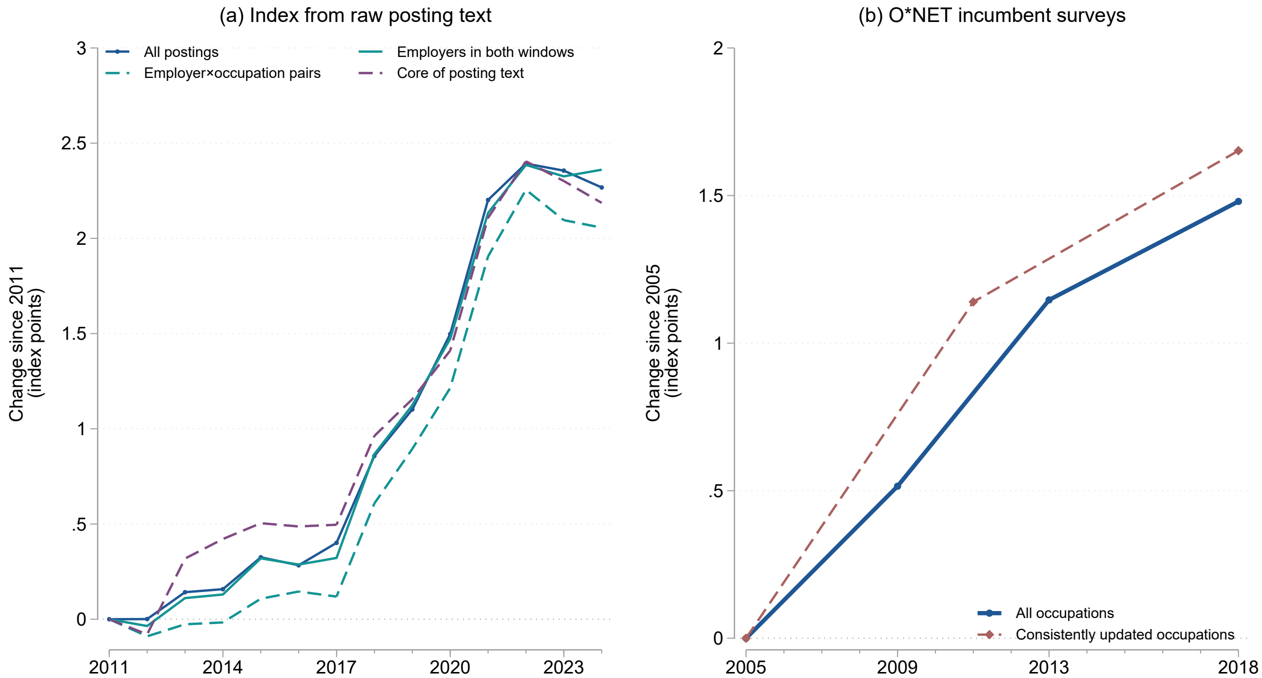
Given the rise in both margins of skill mixing in Figure 4, one may worry that this trend is an artifact of who posts online or of standard openings and closings in posting text. I apply three restrictions to the posting sample to address these concerns. (i) I restrict the sample to employers that post in both the 2010–2012 and 2022–2024 windows, which ensures new types of firms do not drive the trend. (ii) I tighten this restriction to employer-by-occupation pairs present in both windows. (iii) I strip the first and last fifty words of each posting body—where standard openings and closings usually sit—and retain a skill tag only when it appears in the remaining core text.

The additional lines in Figure 4 make the point. On the mixed-skill count in Panel (a), every restricted series climbs steeply with the navy baseline; the employer and employer-by-occupation cuts end 2024 within a few percent of its level, while the core-text series rises strongly from a lower level. On the skill mixing index in Panel (b), the same three cuts are

[24] I omit 2010 from the annual series because Lightcast skill tagging is sparse in its first year. Consequently, the 2010 cross section sits well below the 2011–2024 path.

[25] By comparison, social skill task inputs grew by about 12 percentiles of the 1980 distribution over the 32 years from 1980 to 2012 (Deming, 2017)

Figure 5: The Trend Outside Lightcast's Tagging: Raw Posting Text and O*NET



Notes: This figure plots the skill mixing index outside Lightcast's parsed tags. Panel (a) searches raw posting text for the task keywords of Hershbein and Kahn (2018) and Braxton and Taska (2023), at commuting-zone × four-digit occupation cells, employment-weighted, as the change since 2011. The navy series rises 2.3 points and the unweighted cell mean rises 4.2 points. The panel redraws the three restricted samples of Figure 4. Panel (b) plots the same index from O*NET incumbent surveys as the change since 2005 in the employment-weighted mean raw PCA index, multiplied by 100. The solid line uses all occupations. The dashed line uses the 274 seven-digit occupations resurveyed in 2005, 2011, and 2018.

redrawn and every line rises, with the employer and pair endpoints within about a fifth of the baseline and the core-text series ending above it. Therefore, the rise in skill mixing is not tied to who posts or to standard openings and closings in posting text. The rest of the paper measures degree with the index and uses the count as a detection margin.

Alternative parser. The posting-sample restrictions already leave the rise intact under who posts and after stripping standard openings and closings. However, there may be concern that Lightcast's skill tags artificially create this trend. If the parser were generating broader skill requirements, analyzing skill demand directly from vacancy text would not show the same pattern. Panel (a) of Figure 5 addresses that concern by rebuilding the same mixing index from the fixed keyword lists of Hershbein and Kahn (2018) and Braxton and Taska (2023), which map directly into the four RNR groups—Analytical, Interpersonal, Computer, and Routine—listed in Appendix Table A2. The table also records the posting-level assignment route used here. Beginning with that panel, the body-text index rises 2.3 points from the early

2010s to 2024—about the 50th to the 56th percentile of the 2011 cell distribution—and every plotted restriction series rises with it. The climb is therefore not an artifact of Lightcast's parser.

An independent source. A natural complement to Lightcast is O*NET—the Department of Labor successor to the Dictionary of Occupational Titles, and a primary source for occupational skill requirements in the task literature (Acemoglu and Autor 2011; Yamaguchi 2012; Deming 2017). While the earlier versions of O*NET include legacy ratings from DOT analysts, a shift occurred in 2003 when O*NET began sourcing responses from random samples of workers (job incumbents). For consistent measurement I use only descriptors from questionnaires updated by worker surveys.[26] Following Acemoglu and Autor (2011), I form the same four RNR intensities from those modules; Appendix Table A4 lists the descriptors. The mixing index is the same cosine measure as above. A key challenge when using O*NET comes from employing the longitudinal variation in occupation descriptors. Specifically, while each version of O*NET contains roughly 970 7-digit occupations, an average of 110 occupations undergo updates annually. To capture more granular time patterns, I use 4-year intervals—2005, 2009, 2013, and 2018—ensuring updates to cover over half of the occupations within these intervals.

Panel (b) of Figure 5 demonstrates that the degree of skill mixing has risen substantially and steadily between 2005 and 2018 using O*NET data. The employment-weighted mean of the index rises 1.5 points—about the 50th to the 60th percentile of the 2005 distribution—and gains at every wave. A potential concern of using O*NET data is that the trend could be affected by the inconsistency in occupation updating. Restricting attention to the 274 7-digit occupations that are constantly updated between 2005, 2011, and 2018, thus reflecting a consistent updating of skill requirements among these occupations, gives essentially the same result. The index rises 1.7 points, again about the 60th percentile.

I.D.2 Decomposition

Within versus across occupations. Does the rise come from occupations changing what they demand, or from employment sliding toward already-mixed occupations? The left block of Table 1 presents shift-share decomposition of each series into *within*-occupation change and *across*-occupation reallocation. Growth in every series is driven by changes within occupations.

[26] Specifically, I use descriptors from the Work Context, Work Activities, Knowledge, and Skills questionnaires. After 2003, O*NET still contains responses from job analysts for questionnaires that have small sample sizes from workers. I abstract from those questionnaires in this paper.

Table 1: The Rise in Skill Mixing: Decomposition and Annual Trends

	Full-period change: within/across decomposition			Annual trend within local labor markets (coefficient on year)			
	Total	Within	Across	No FE (1)	CZ + occ. FE (2)	CZ × occ. FE (3)	+ worker controls (4)
No. mixed skills (Figure 4a)[†]	56.47	53.09	3.38	0.139***	0.142***	0.142***	0.121***
				(0.002)	(0.002)	(0.002)	(0.005)
Skill mixing index (Figure 4b)	5.13	5.17	−0.04	0.380***	0.578***	0.587***	0.317***
				(0.059)	(0.061)	(0.062)	(0.074)
Keyword-search index, posting text (Figure 5a)	2.27	2.58	−0.31	0.218***	0.238***	0.242***	0.217***
				(0.016)	(0.016)	(0.017)	(0.021)
O*NET survey index, 2005–2018 (Figure 5b)[‡]	1.48	1.33	0.15	0.137***	0.138***	0.137***	0.131***
				(0.030)	(0.029)	(0.042)	(0.039)

Notes: This table reports the four series of Figures 4 and 5. The left block decomposes the full-period change into within-occupation and across-occupation parts, $\Delta T = \sum_j \bar{E}_j \, \Delta\alpha_j + \sum_j \Delta E_j \, \bar{\alpha}_j$, each row in its plotted units. The skill-mixing-index and keyword rows use 2011–2024 commuting-zone × four-digit ISCO-08 cells on the 50–100 cosine scale, employment-weighted. In 2011-anchored ranks the same splits are 9.49 (9.50 within) and 6.06 (6.40 within). The tag-count row uses national four-digit ISCO shares. The O*NET row decomposes the 2005–2018 rise (1.65 within on the 274 resurveyed codes); on the ACS occ1990dd panel the rise is 1.91 points (10.12 ranks). The right block is the OLS coefficient on year. Posting rows are commuting-zone × four-digit ISCO × year, 2011–2024: (1) no FE; (2) commuting-zone and two-digit ISCO FE; (3) their interaction; (4) plus lagged commuting-zone-by-year worker controls (Autor, Dorn, and Hanson, 2013). [†]Left block: mixed tags per 100 postings (Figure 4a units). Right block: commuting-zone × occupation cell regression of $\log(1 + \text{count})$ on year (column (4): 0.121). [‡]O*NET trend columns use the ACS occupation × demographic cell panel at waves 2005, 2009, 2013, and 2018: (1) no FE; (2) sex × education; (3) sex × education × industry × occupation; (4) plus experience controls. Standard errors cluster by commuting zone for posting rows and by occupation (338) for O*NET. 2024 weights use the 2023 ACS. Share-weighted tag index, keyword-on-tags, and mixed-posting-share variants are 4.46, 1.80, and 18.94 percentage points. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

For the mixing index, the entire 5.1-point increase comes from within-occupation changes, with 5.2 points within and almost zero across occupations. In the mixed-skill count, 53.1 out of the 56.5-tag rise per 100 postings happens within occupations, while employment reallocation adds 3.4. The survey benchmark shows a similar pattern: of the O*NET row's 1.48-point increase from 2005 to 2018, 1.33 points are within occupations and 0.15 across them. In summary, changes within occupations explain the rise, which matches the job-content shifts Atalay et al. (2020) found within occupations from 1950 to 2000.

Within local labor markets. The right block of Table 1 asks whether mixing rises within local occupational labor markets once worker composition is held fixed. For each posting-based measure I fit a linear trend in the commuting-zone × four-digit ISCO-08 × year panel over 2011–2024. Column (1) is the raw trend; columns (2)–(3) add commuting-zone and two-digit occupation fixed effects, then their interaction, so identification comes from year-to-year change within a local-occupation cell; column (4) adds lagged commuting-zone-by-year controls following Autor, Dorn, and Hanson (2013). Standard errors are clustered by commuting zone. Since O*NET measures are national by occupation and carry no geography, the O*NET columns run the analogous composition check on the ACS occupation $\times$ demographic cell panel at the four survey waves, absorbing sex–education and industry–occupation cells in place of the CZ ladder.

Every measure rises under every specification. Beginning with the mixed-skill count, the annual gain in $\log(1 + \text{count})$ is 0.12–0.14 a year and barely moves when local and occupation fixed effects or worker controls are added. On the skill mixing index the raw slope is 0.38 points a year. The full controls reduces it slightly to 0.32. The keyword-search series has the same shape. The O*NET index gains 0.14 without controls and 0.13 with them. Therefore, worker composition does not explain the skill mixing growth.

To summarize, occupational skill demand has become more mixed. Nearly half of what an occupation lists in 2024 was absent from its own 2010 postings, and almost all of that newly observed content was already demanded elsewhere—predominantly general-purpose, non-routine content, originating in knowledge-intensive occupations and entering lower-skill ones. The skill mixing index rises from the early 2010s under every restriction of the posting sample and in independent surveys of incumbents; the rise is within occupations and within local labor markets, with little role for employment reallocation or worker composition. A within-cell change in job content points to a demand-side account, and the next section asks which technology, if any, produces it.

II Measuring Breakthrough GPTs

What drives the rise in skill mixing? I examine breakthrough general-purpose technologies—innovations that are highly novel and broadly applicable. Unlike incremental innovations that refine existing processes, GPTs generate "innovational complementarities" across a broad range of economic activities (Bresnahan and Trajtenberg 1995; Helpman and Trajtenberg 1994) and induce organizational changes (Bresnahan et al. 1996). This section builds a measure of occupational exposure to them from the universe of U.S. patents, linking patent text to the technology descriptors of each occupation.

The exposure measure must capture three elements. The innovation must represent a genuine departure, its influence must be broad, and it must reach the occupations whose technological requirements it changes. I build all three from the U.S. patent record in the USPTO's PatentsView database. I measure novelty using the text-based breakthrough measures of Kelly et al. (2021) and breadth using a citation-based generality index. I then map the patents that meet both criteria to occupations by comparing each patent's abstract with O*NET descriptions of the technologies used in that occupation.

II.A Identifying Breakthrough GPTs

Breakthrough novelty. I use the text-based breakthrough measure of Kelly et al. (2021). It compares a patent's abstract with earlier and later patents. Low similarity to earlier patents, combined with high similarity to later ones, identifies patents that depart from an existing line of work and are subsequently followed. Following Kelly et al. (2021), I classify patents above the 90th percentile of the published score as breakthroughs. The measure captures novelty instead of generality, as a patent can open a new line of work without influencing multiple fields.

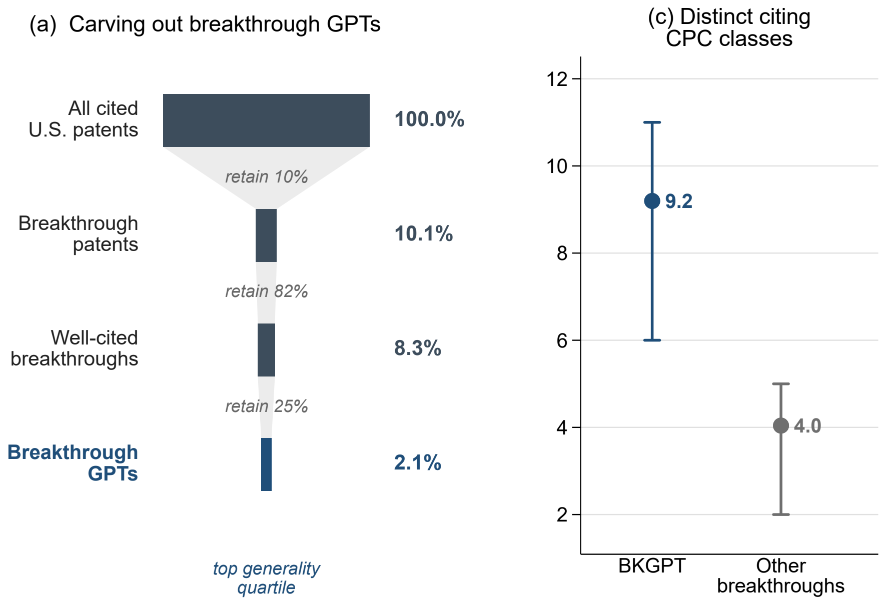
Generality. A GPT's signature is the breadth of its downstream influence, which I measure by how widely a patent's forward citations disperse across technological fields. For a patent p, the Hall, Jaffe, and Trajtenberg (2001) generality index is $Gen_p = 1 - \sum_c s_{pc}^2$, where s_{pc} denotes the share of forward citations received from patents in technology class c (based on the Cooperative Patent Classification 2-digit level). A higher index indicates that the innovation influences a broader range of subsequent technological fields, the key feature of a GPT. I define a patent as having "High Generality" ($HighGen_p = 1$) if its index falls in the top quartile of the generality distribution among breakthrough patents.[27] A patent that satisfies both criteria simultaneously is a *Breakthrough GPT*:

$$\text{Breakthrough GPT}_p = \mathbb{1}\{HighGen_p = 1 \wedge BK_p = 1\}.$$

The joint cut is deliberately narrow. Figure 6 shows that about a tenth of cited patents are breakthroughs and only about two percent are breakthrough GPTs. The ones that pass are cited across roughly twice as many 2-digit CPC classes as other well-cited breakthroughs (9.2 versus 4.0), even as their citation totals are only modestly larger (107 versus 62). Breadth, not the raw amount of attention, is what the definition isolates. A single patent makes the point. An airborne fulfillment center that dispatches delivery drones, filed in 2014, draws its 456 forward citations from 25 distinct 2-digit CPC classes. A pharmaceutical compound or a surgical device, however heavily cited, almost never reaches that many fields.

[27] Citation classes are taken at the 2-digit CPC level and the index is computed over all forward citations a patent receives. The quartile cut is taken within the pool of breakthrough patents with at least five forward citations, so that generality is well measured.

Figure 6: Defining Breakthrough GPTs and Their Distinctive Breadth



Notes: This figure plots how breakthrough GPTs are selected from cited U.S. patents. Panel (a) reports each stage as a share of all cited patents. A patent is a breakthrough if its Kelly et al. (2021) text-based score is in the top decile. Among breakthroughs with at least five forward citations, Breakthrough GPTs are the top quartile of the Hall, Jaffe, and Trajtenberg (2001) generality index. Panel (c) compares Breakthrough GPTs with the remaining well-cited breakthroughs by the number of distinct 2-digit CPC classes among forward citations. Markers are class means and whiskers span the 25th–75th percentile range.

Panel A of Table 2 lists high-generality breakthroughs—advanced materials, energy conversion, logistics platforms, and autonomous systems—whose citations span about twenty fields. Panel B lists breakthroughs that are similarly cited but specialized—pharmaceutical compounds, prosthetic and surgical devices, and point-to-point telecommunications hardware. The patents in the two panels are similarly novel and similarly cited. What separates them is the range of technologies that build on them, which is the distinction the exposure measure is designed to capture.

II.B From Patent Text to Occupation Exposure

A patent affects an occupation only if it describes a technology used in that occupation. I measure this link by comparing the text of breakthrough-GPT patents with O*NET Technology Skills descriptions.

For each patent–occupation pair, I compare the patent abstract with the occupation's technology descriptors and retain close matches.[28] The cutoff is deliberately high: only pairs

[28]The embedding model is a transformer rather than a static word-vector model such as Continuous Bag of Words (CBOW). CBOW learns one fixed vector per word by predicting a target word from the words

Table 2: Breakthrough GPTs Versus Specialized Breakthroughs

	Patent title	Filed	Generality	Citing CPC classes	Forward citations
Panel A. Breakthrough GPTs (top generality quartile)					
1	Autonomous data machines and systems	2015	0.935	21	296
2	Carbon nanotube composite and method for fabricating the same	2006	0.934	20	80
3	Nanomesh article and method of using the same for purifying fluids	2005	0.931	39	241
4	Docking process for recharging an autonomous mobile device	2011	0.931	18	441
5	Airborne fulfillment center utilizing unmanned aerial vehicles for item delivery	2014	0.930	25	456
6	Air power energy transformation to electrical energy for hybrid electric vehicle applications	2009	0.929	29	129
7	Managing autonomous machines across multiple areas	2010	0.929	18	370
8	Food preparation system	2005	0.929	22	98
Panel B. Specialized breakthroughs (bottom generality quartile)					
1	Bupropion as a modulator of drug activity	2016	0.000	1	99
2	Compositions and methods for increasing the metabolic lifetime of dextromethorphan	2016	0.000	1	94
3	Prostheses for replacement of natural facet joints	2002	0.036	3	325
4	Terminal adapter interfacing an ATM network with an STM network	1994	0.054	4	219
5	Intervertebral endoprosthesis with a toothed connection plate	2000	0.068	4	286
6	Data packet radio service with enhanced mobility management	1998	0.089	4	152
7	Variable length electrodes for delivery of irrigated ablation	2000	0.162	7	269
8	Bifurcated endoluminal prosthesis	1997	0.283	9	444

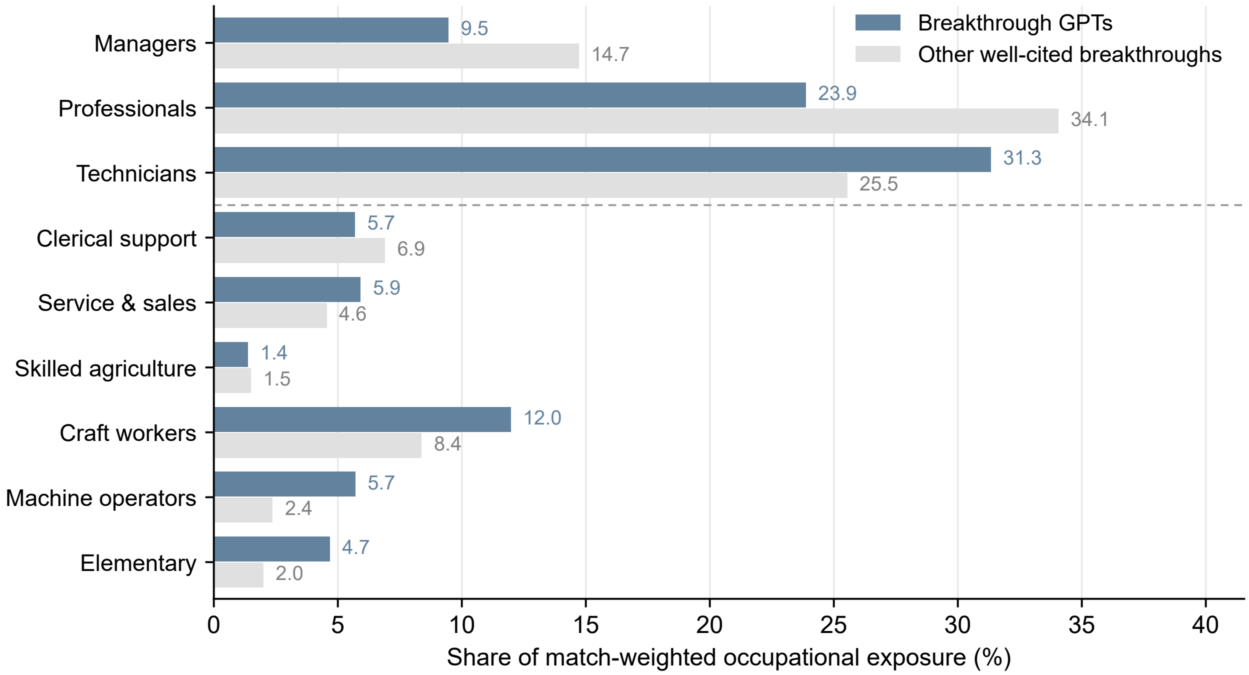
Notes: This table lists influential breakthrough patents (top decile of the Kelly et al. 2021 text-based breakthrough score, with at least 25 forward citations), ranked by the Hall, Jaffe, and Trajtenberg (2001) generality index $Gen_p = 1 - \sum_c s_{pc}^2$, where s_{pc} is the share of forward citations from 2-digit CPC class c. Panel A shows patents in the top generality quartile (Breakthrough GPTs). Panel B shows breakthrough patents of comparable citation impact in the bottom generality quartile.

above a cosine similarity of $\tau_0 = 0.45$ enter the measure.[29] I assign greater importance to closer matches and to patents whose citations span multiple fields by weighting them by excess similarity and citation generality. Thus, a broadly useful patent contributes most

immediately around it; a transformer represents a full text span in context, so it can register that two documents describe the same technology in different vocabulary. Embeddings are 3,072-dimensional, from OpenAI's `text-embedding-3-large`.

[29]Fewer than 20 percent of breakthrough patents clear this cutoff on their best descriptor.

Figure 7: The Occupational Footprint of Breakthrough GPTs



Notes: This figure plots match-weighted occupational exposure of breakthrough-GPT patents and other well-cited breakthroughs by ISCO-1 major group. Matching uses the same embedding-based patent–descriptor matches, excess-similarity weights, and crosswalk chain as the exposure measure (filing years 1987–2015). Breakthrough classification follows Kelly et al. (2021). High generality is the top quartile of the Hall, Jaffe, and Trajtenberg (2001) index among well-cited breakthroughs. Section II details the matching rule.

where it closely matches an occupation's existing toolkit. The cutoff and weights were fixed before estimating any labor-market effects; Appendix Table A6 reports coverage under the continuous rule at alternative cosine thresholds τ_0, and Table A7 reports the corresponding coefficient robustness.

Figure 7 asks where these matched patents land across ISCO-1 occupation groups. Both breakthrough GPTs and other well-cited breakthroughs concentrate on the professional and technical core. What differs is the tail: roughly 35% of the breakthrough-GPT footprint falls on jobs *outside* the professional-technical-managerial core, against 26% for other well-cited breakthroughs, and the gap is two to three times as large among operators and elementary occupations. Patents that are broad in the citation record are also broad in the occupations they reach.

Matching alone does not say when a technology matters for jobs. Because technologies diffuse over many years, I measure exposure as a depreciated patent stock rather than an annual flow. Knowledge stocks in R&D are commonly built with geometric depreciation (Griliches 1979; Hall, Jaffe, and Trajtenberg 2005; Bloom, Schankerman, and Van Reenen

2013), and recent labor applications do the same for installed robots and patents (Acemoglu and Restrepo 2020; Autor et al. 2024). The reason is simple: new technologies diffuse over many years (Comin and Hobijn 2010). I therefore cumulate the matched breakthrough-GPT patent-equivalents into an annual occupation count g_{ot}, and apply an annual depreciation rate of $\delta = 0.10$.[30] This stock is the primary measure of exposure throughout Section III. Alternative depreciation rates and flow-based variants are discussed therein.

III The Effects of Breakthrough GPTs

This section asks whether breakthrough GPTs are the force behind the rise in skill mixing, and what their arrival does to wages and employment. I combine the exposure measure of Section II with Lightcast postings and American Community Survey microdata, in a panel of commuting zones and demographic cells over 2010–2024. A shift-share design, instrumented with California-based innovation, identifies the effect of exposure. The first subsection sets out that design—the panel, the instrument, and the specifications the tables report. The next two estimate the response of skill mixing, wages, and employment. The fourth races breakthrough-GPT exposure against routine task automation, the leading alternative account of how technology reshapes skill demand. The last subsection examines which skills arrive and which workers bear the incidence.

III.A Shift-Share Identification Strategy

The empirical strategy is a shift-share design in the recent labor-economics tradition (Acemoglu and Restrepo 2020; Borusyak, Hull, and Jaravel 2022; Borusyak and Hull 2023; Adão, Kolesár, and Morales 2019). The comparison is straightforward. A demographic group whose members in 2010 worked disproportionately in occupations later reached by breakthrough-GPT shocks experiences a differentially larger aggregate exposure—a Bartik (1991) shift-share design in which the "shares" are baseline occupation weights within the demographic cells, and the "shocks" are national. The identifying variation is national and technological. California originates a disproportionate share of frontier GPT patenting, while the rest of the country is exposed to that innovation. Following Borusyak, Hull, and Jaravel (2022) and Borusyak and Hull (2023), I recenter the occupation-level shock through year demeaning, so that consistency requires only that breakthrough-GPT innovation be assigned across occupations as good as

[30] Because the Kelly et al. (2021) breakthrough score requires forward text similarity, it is unavailable for recent filing vintages; usable breakthrough-GPT inflows end around 2015, and from 2016 the stock evolves only through geometric depreciation.

randomly. The rest of this subsection lays out the panel, the California instrument, and the specification ladder the tables follow.

Panel and exposure. The unit of analysis is a commuting zone $\times$ demographic cell $\times$ year, where a demographic cell k = (sex $\times$ education $\times$ potential experience) is defined by two sexes, five education bins, and five experience categories, observed annually over 2010–2024.[31] All occupation coding uses the ISCO-08 taxonomy at the 2-digit level. Two data features motivate this choice. First, Lightcast tags each posting to ISCO-08 natively. Furthermore, the two-digit level is sufficiently granular to capture distinct breakthrough-GPT shocks, but coarse enough to construct a clean shift-share weight without generating empty demographic cells. For the ACS, I first map Census occupation codes into SOC-2010, and then map SOC-2010 into ISCO-08 using the crosswalk published online by Institute for Structural Research (2016).[32] The shift-share exposure of cell (c, k) is

$$B_{ckt} \;=\; \sum_{o} s_{cko}^{2010}\, g_{ot}, \tag{2}$$

where s_{cko}^{2010} is the employment share of ISCO-2 occupation o within cell (c, k), fixed at its 2010 value to pre-date the main shock period, and g_{ot} is national breakthrough-GPT patent intensity for occupation o in year t.[33] The estimating equation is

$$Y_{ckt} \;=\; \beta\, B_{ckt} + X_{ct}'\Gamma + \delta_c + \delta_{kt} + \delta_{s(c)t} + \varepsilon_{ckt}, \tag{3}$$

where Y_{ckt} is the outcome and X_{ct} is the commuting-zone covariate vector—the standard set of pre-shock local characteristics (Census manufacturing and routine employment shares, offshorability, demographic composition, Chinese import exposure, and local wage and employment indicators).[34] The panel further includes commuting-zone (δ_c) fixed effects, demographic-cell$\times$year (δ_{kt}) fixed effects that absorb national, time-varying shocks common

[31] Working at the demographic-cell level (rather than at the occupation level) allows the wage and employment outcomes to be defined for well-populated ACS cells and lets heterogeneity be traced across worker education and across the skill content of the requirements that arrive (Section III.E); the occupational dimension enters through the baseline share weights of the shift-share.

[32] The implemented chain is Census OCC to SOC-2010, then to ISCO-08. The code joins Census occupation codes to the OES-2010 OCC–SOC bridge, uses OES employment weights to allocate multi-code matches, applies the IBS SOC-2010–ISCO-08 crosswalk, and aggregates to ISCO-2. The normalized splits are additive, so the mapping conserves cell totals.

[33] Fixing shares at 2010 rather than at a rolling average sharpens the predetermined-share interpretation of the design (Goldsmith-Pinkham, Sorkin, and Swift, 2020): the variation in B_{ckt} over time within a cell comes entirely from the national shocks g_{ot}, not from contemporaneous reallocation across occupations.

[34] These covariates control for the pre-existing local trajectories most likely to correlate with both baseline occupation mix and later labor-market outcomes. Observations are weighted by 2010 cell size.

to a sex-education-experience group, and state×year ($\delta_{s(c)t}$) fixed effects that absorb regional business cycles and policy. Standard errors are two-way clustered by commuting zone and state.

Identification and recentering. Equation (3) estimated by OLS requires that the national shocks g_{ot} be quasi-randomly assigned across occupations conditional on the shares. This may fail if occupations that systematically attract more patent activity, breakthrough or not, also trend differentially in skill demand. I therefore instrument B_{ckt} with a recentered shift-share instrument built from California-based patent assignees alone, following Borusyak and Hull (2023) and Borusyak, Hull, and Jaravel (2022).

The instrument is a *California-recentered* shock that isolates innovation originating at the technological frontier. California produces a disproportionate share of breakthrough GPT patents, while other regions are exposed to these advances.[35] Within each year I subtract the year-mean California-only shock across all ISCO-2 occupations before aggregating with the 2010 occupation shares,

$$B_{ckt}^{\mathrm{CA}} \;=\; \sum_{o} s_{cko}^{2010}\left(g_{ot}^{\mathrm{CA}} - \bar{g}_{t}^{\mathrm{CA}}\right), \qquad \bar{g}_{t}^{\mathrm{CA}} = \frac{1}{|\mathcal{O}|}\sum_{o'\in\mathcal{O}} g_{o't}^{\mathrm{CA}}, \tag{4}$$

where g_{ot}^{CA} uses only patents whose assignees are based in California and $\mathcal{O}$ indexes ISCO-2 occupations. This isolates a supply-side source of innovation—one major patenting cluster—that is plausibly exogenous to labor-demand conditions elsewhere.[36] The design rests on one transparent exclusion restriction: California-assignee innovation is uncorrelated with non-California labor-demand shocks, conditional on the fixed effects and controls in equation (3).

Equation (4) subtracts the year-mean shock to apply the recentering approach of Borusyak and Hull (2023) under unconditional exchangeability, which assumes that within a year, breakthrough-GPT intensity is assigned as good as randomly across all ISCO-2 occupations. Year-demeaning removes the component of the California shock common to all occupations

[35] The exclusion logic follows the robots-and-jobs strategy of Acemoglu and Restrepo (2020), who instrument U.S. local robot exposure with robot adoption in European industries and build exposure from 1990 industry employment shares—so that common industry-level adoption abroad captures advances at the technological frontier while remaining less likely to reflect U.S.-specific labor-demand shocks. Later work uses the same European frontier to isolate automation driven by international technological advances (Acemoglu and Restrepo 2022,?). The present design is the occupational analogue. California-based breakthrough-GPT innovation replaces their external frontier, an occupation-level shock replaces their industry-level one, and 2010 ISCO-2 occupation shares replace their 1990 industry shares.

[36] As a leave-out safeguard the cumulative instrument also excludes the single highest-leverage ISCO-2 occupation from the share basis, and the exposure and instrument are both scaled to the same standard-deviation footing (see the discussion of units below Table 3).

in a given year; it does not by itself purge confounders that operate at a coarser occupational level (Goldsmith-Pinkham, Sorkin, and Swift, 2020; Borusyak and Hull, 2023). One may still worry that patenting concentrates in professional and technical jobs, so the identifying contrast is mostly between broad occupation groups. Section III.C discusses alternative recentering.

The specification ladder. Table 3 columns (1)–(2) report OLS without and with the commuting-zone covariates. Column (3) reports 2SLS using the raw California leave-out instrument without those covariates, and column (4) the same specification with the year-demeaned instrument, so the move from (3) to (4) isolates recentering. Column (5) adds the covariate vector to column (4) and is the baseline under unconditional exchangeability, so the move from (4) to (5) isolates the controls. Column (6) retains the baseline instrument and controls but excludes the five largest patent-producing states (CA, NY, TX, WA, MI).

III.B Baseline Results

Table 3, column (5), presents the baseline estimates for the effect of a one-standard-deviation increase in exposure.[37] A one-standard-deviation increase raises the skill mixing index by 1.89 points on the 50–100 scale (SE 0.21, $t = 9.2$). The Anderson–Rubin 95 percent interval strictly excludes zero. To place this estimate in the context of the trend from Section I (Figure 4), a one-standard-deviation rise accounts for roughly one third of the 5.1-point secular increase in the index.

The detection margins of skill mixing move with the mixing index. The *extensive* margin—the probability that a posting lists any mixed skill at all—rises by 3.2 percentage points per standard deviation (SE 0.2). The *intensive* margin—the log number of mixed skills among postings that already list one, $\ln(n)$ with $n \geq 1$—rises by 37.1 log points (SE 3.2).[38] Taken together, the three rows of Panel A say one thing: exposure raises the balance of skills within postings, the breadth of skills demanded, and the probability that a posting spans skills at all.

Panel B examines labor market outcomes. A one-standard-deviation increase in exposure increases hourly wages by 14.0 log points (SE 1.0). The employment rate remains unchanged,

[37] One standard deviation corresponds to 19.0 patent-units of exposure, weighted by 2010 employment. A raw patent-unit represents a similarity- and citation-weighted count with no natural cardinal meaning. Dividing any entry by 19.0 recovers the per-patent-unit coefficient. The scaling does not affect t-statistics, F statistics, or Anderson–Rubin coverage.

[38] A log point is one hundredth of a log unit, so the 0.371 coefficient of Table 3 is 37.1 log points. The intensive margin is $\ln(n)$ among postings with $n \geq 1$ mixed skills (Chen and Roth, 2024). All three posting margins are aggregated to commuting-zone $\times$ demographic-cell $\times$ year with 2010 occupation shares, matching the skill mixing index.

Table 3: Effect of Breakthrough-GPT Exposure on Skill Mixing and Labor Markets

	OLS		Just-identified 2SLS			
	bare	full	CA stock (raw)	year-demeaned	**baseline** year-demeaned	drop top-5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Skill mixing in job postings						
Skill mixing index (×100)	1.51***	1.48***	2.51***	2.15***	1.89***	1.80***
	[0.20]	[0.19]	[0.27]	[0.22]	[0.21]	[0.17]
Prob. any mixed skill (extensive)	0.022***	0.021***	0.041***	0.036***	0.032***	0.030***
	[0.002]	[0.002]	[0.003]	[0.002]	[0.002]	[0.002]
Log # mixed skills (intensive)	0.139***	0.137***	0.468***	0.368***	0.371***	0.324***
	[0.012]	[0.012]	[0.049]	[0.033]	[0.032]	[0.024]
Panel B. Local labor-market outcomes						
Log hourly wage	0.071***	0.072***	0.169***	0.140***	0.140***	0.124***
	[0.003]	[0.003]	[0.014]	[0.010]	[0.010]	[0.006]
Employment rate (ACS)	0.098*	0.106*	0.003	0.029	0.026	0.057
	[0.054]	[0.057]	[0.130]	[0.092]	[0.093]	[0.074]
First-stage coefficient	—	—	2.10	2.96	2.96	3.05
Effective F (shock-level cluster)	—	—	7.8	58.0	58.6	76.0
Shock clusters (ISCO-2)	—	—	38	38	38	38
Observations	531,120	481,073	531,035	531,120	481,073	398,934
CZ covariates	—	Yes	—	—	Yes	Yes
CZ, cell×year, state×year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the effect of a one-standard-deviation increase in breakthrough-GPT exposure, equal to 19.0 patent-units (2010-employment-weighted). Panel A, first row, is the skill mixing index of Section I.C on the 50–100 scale. Columns (1)–(2) are OLS without and with commuting-zone covariates. Columns (3)–(4) are just-identified 2SLS without those covariates: column (3) uses the raw California leave-out instrument; column (4) year-demeans shocks across ISCO-2 occupations (equation (4)) before aggregating. Column (5) adds commuting-zone covariates to column (4) and is the baseline. Column (6) drops the top-5 patent states (CA, NY, TX, WA, MI). All specifications include commuting-zone, demographic-cell-by-year, and state-by-year fixed effects, 2010-employment weights, and two-way clustering by commuting zone and state. The ACS employment-rate sample is 411,784 in column (5); column (3) loses 85 cells that lack the raw leave-out instrument. The Anderson–Rubin 95 percent interval for the skill mixing index in column (5) is $[1.44, 2.27]$. The effective F clusters the first stage at the shock level (38 ISCO-2 clusters). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

with an estimate of 0.026 (SE 0.093), indicating no significant job loss in exposed cells as wages rise. Therefore, these results indicate higher wages without employment displacement.

Other Specifications. The California instrument uses a geographically separate patent pool to isolate shared occupational variation, and the resulting IV estimates exceed OLS for both mixing and wages. Demeaning shocks by year lowers the mixing estimate without covariates from 2.51 to 2.15 and the wage estimate from 0.169 to 0.140 in columns (3) and (4). Demeaning strengthens the design by concentrating identification at the shock level. Adding the covariates yields the baseline mixing estimate of 1.89 and leaves the wage estimate at 0.140.

One might wonder that whether the baseline results merely reflect the local labor market dynamics of California or other high-patent states. Column (6) drops the five largest patent-producing states (California, New York, Texas, Washington, and Michigan) from the estimation sample but retains the California-built instrument. The estimates remain essentially unchanged. For example, the mixing index coefficient shifts from 1.89 to 1.80, and the wage coefficient shifts from 0.140 to 0.124. Because the estimates persist after excluding these major suppliers, the internal labor market of any single high-patent state cannot drive the results.

Inference. Clustering on ISCO-2 occupation shocks yields effective F statistics between 58 and 76 in the year-demeaned columns, exceeding the Stock-Yogo thresholds for designs with a modest number of independent shocks. Anderson–Rubin intervals, which remain valid under weak instruments (Appendix Table A8), exclude zero for all posting-based mixing margins and wage outcomes, but the employment interval includes zero. There is a potential concern that two-way cell-level clustering may overstate the independence of these shocks. Appendix Table A9 presents Adao–Kolesár–Mikusheva standard errors for the same column (5) coefficients; mixing and wage effects remain significant, while the employment interval includes zero under both standard error methods.

The exclusion restriction. The identifying assumption is that California-assignee innovation is uncorrelated with non-California labor-demand shocks, conditional on the fixed effects and covariates. I run three more checks. First, when I drop the top five patent-producing states, the main results stay the same (see Table 3, column 6). This suggests the effect comes from diffusion, not just from any single state with many patents. Second, the pre-period diagnostics in Appendix Table A10 show that skill-mixing, wages, and employment rates did not move differently before exposure. The pre-shock placebos are flat, so the main effects are not explained by earlier trends before breakthrough-GPT spread. Third, someone might worry that jobs later affected by the shock were already changing in skill content. But, after accounting for the 2010 mixing level (which the fixed-effect stack controls for), the change from 2011 to 2013 does not predict later exposure ($p = 0.46$, Appendix Table A10, Panel B).

III.C Additional Robustness

Alternative recentering. The baseline recenters the California shocks within year. This removes the common annual component, but it still compares occupations across broad ISCO-1 groups. I therefore recenter within ISCO-1 group and year, holding the exposure measure, baseline shares, sample, controls, and fixed effects unchanged (Appendix A.8). The skill mixing index falls modestly, from 1.89 to 1.48 points per standard deviation, and all three mixing responses remain significant under within-group randomization inference. The employment estimate remains near zero. The wage coefficient falls from 14.0 to 5.6 log points. Thus, the mixing result is robust to using only within-group variation, while the magnitude of the wage response is more sensitive to comparisons across broad occupation groups.

Shock leverage and weight. I next ask whether any broad occupation group drives the estimates. In the leave-one-ISCO-out re-estimations (Appendix Table A12), the skill mixing index and the posting detection margins remain positive in every drop, the wage coefficient is significant at the one-percent level in all nine drops, and the employment-rate coefficient is imprecise throughout. The highest weight is placed on Technicians and Associate Professionals, and at the level where the shock is aggregated, HHI is 0.031.[39] The central mixing results therefore do not depend on retaining any one occupation group.

Unclassified skills. Finally, the mixing result does not depend on excluding skill tags that cannot be assigned to the four content categories. Treating these tags as a fifth category raises the coefficient from 1.89 to 2.19 points per standard deviation (SE 0.19) (Appendix Table A13). Excluding unclassified skills therefore attenuates, rather than generates, the estimated response.

III.D Distinguishing Breakthrough GPTs from Routine-Task Automation

Routine-biased technical change is the leading alternative account of how technology reshapes skill demand. To compare the two, I construct the Autor and Dorn (2013) routine-task intensity at the ISCO-2 level, project it onto the same 2010 employment shares, similar to the contruction of the breakthrough-GPT shock, and interact that predetermined exposure with a linear trend:

$$Y_{ckt} = \beta_G B^{\text{GPT}}_{ckt} + \beta_R \left(R^{\text{RTI}}_{ck} \times t\right) + X'_{ct}\Gamma + \delta_c + \delta_{kt} + \delta_{s(c)t} + \varepsilon_{ckt}, \tag{5}$$

where B^{GPT}_{ckt} is breakthrough-GPT exposure and R^{RTI}_{ck} is the cell's 2010 routine-task share. The two coefficients are partial effects: β_G holds the routine-task share–trend exposure fixed, while β_R holds breakthrough-GPT exposure fixed. Breakthrough-GPT is instrumented with the California-recentered shift-share counterpart, while $R^{\text{RTI}}_{ck} \times t$ enters as a predetermined control. The race retains the baseline controls, fixed effects, 2010 employment weights, and

[39]Because the regression uses time variation and BKGPT exposure is built from annual patent stocks by occupation, at that aggregation the weighted just-identified estimates reproduce the headline IV and the HHI is 0.031.

Figure 8: Breakthrough GPTs and Routine-Task Automation

Notes: This figure plots the estimates of Table 4. Breakthrough-GPT exposure is instrumented with its California-recentered leave-one-share-out stock. The Autor and Dorn (2013) routine-task measure enters as a predetermined control. Each bar is per one standard deviation of its own exposure. Whiskers are 95 percent confidence intervals, clustered by commuting zone and state (47 state clusters). The count is $\ln(n)$ among postings with $n \geq 1$ mixed skills. The employment panels are the ACS employment rate and the log ACS employment level of the cell ($\times 100$), both for all workers.

two-way clustering by commuting zone and state. Once the fixed effects and covariates are absorbed, the residualised exposures correlate at -0.05.

Figure 8 and Table 4 report the resulting estimates. Including the routine-task competitor leaves the breakthrough-GPT estimates largely unchanged. The mixing coefficient shifts only from 1.89 to 1.91 per SD, and the wage coefficient moves from 0.140 to 0.143. This stability indicates that the two exposures do not compete for the same variation. The routine-task coefficients reflect the standard routine-biased narrative on the pay and extensive content margins. Routine-task exposure yields -0.017 per SD on the wage and -0.004 on the any-mixed probability. For the mixing index itself, the routine-task loading is negative and statistically insignificant (-0.14 per SD, $t = -1.6$). The intensive mixed-skill count loads positively ($+0.045$) but remains an order of magnitude below the breakthrough-GPT effect. A formal test of equality between the per-SD effects rejects the null at $p < 0.001$ for every posting and wage row in Table 4. Breakthrough GPTs pull skills into jobs and raise wages. In

Table 4: Breakthrough GPTs and Automation Exposure

	No competitor	Independent automation measure		
Dep. variable	BK-GPT	BK-GPT	Routine-task	Wald p
Skill mixing index ($\times$100)	1.89***	1.91***	-0.14	<0.001
	[0.21]	[0.20]	[0.09]	
Any mixed skill (0/1)	0.032***	0.032***	-0.004***	<0.001
	[0.002]	[0.002]	[0.001]	
Log # mixed skills (intensive)	0.371***	0.365***	0.045***	<0.001
	[0.032]	[0.031]	[0.009]	
Log hourly wage	0.140***	0.143***	-0.017***	<0.001
	[0.010]	[0.009]	[0.004]	
Employment rate (ACS)	0.026	0.033	-0.039	0.48
	[0.093]	[0.093]	[0.033]	
Log employment level (ACS, $\times$100)	0.33***	0.34***	-0.05***	<0.001
	[0.02]	[0.02]	[0.01]	
Effective F (shock-level cluster)	58.6	68.3		

Notes: This table reports 2SLS for breakthrough-GPT exposure, instrumented with its California-recentered leave-one-share-out stock. The Autor and Dorn (2013) routine-task intensity, interacted with a linear trend, enters as a predetermined control. Each cell is per one standard deviation of its own exposure (19.0 patent-units for breakthrough-GPT exposure and 2.58 for the routine-task measure). Wald p tests equality of the per-SD effects. The count is $\ln(n)$ among postings with $n \geq 1$ mixed skills. Employment rows are the ACS employment rate and the log ACS employment level ($\times$100). The panel, weights, fixed effects, and covariates match Table 3, column (5). $N = 481{,}073$ for skill and wage rows and 411,784 for the employment rate. Standard errors are clustered by commuting zone and state (47 state clusters). The effective F clusters the first stage at the shock level (38 clusters). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

contrast, routine-task exposure lowers wages, depresses the any-mixed margin, and produces no significant effect on the mixing index.

The last two panels of Figure 8 turn to employment. On the employment *rate*, which measures whether a cell's people work, both coefficients are indistinguishable from zero, and the equality test that rejects on every other row of Table 4 fails to reject here ($p = 0.48$). On the log employment *level*, which measures how much employment a cell carries, the two diverge. Routine-task exposure lowers employment by about 0.05 percent per SD, and breakthrough-GPT exposure raises it by about 0.34 percent.[40] Neither exposure changes the share of people who work. That pattern on the rate is consistent with the reallocation account in Autor and Dorn (2013), in which routine decline shifts workers across occupations without changing who is employed. Exposed cells simply end up with less employment under the routine-task measure and more under breakthrough GPTs.

Taking stock, breakthrough-GPT exposure and the Autor–Dorn routine-task measure are not the same technology under different labels. Breakthrough GPTs raise the skill mixing, wages, and the employment level of exposed cells. Routine-task exposure lowers wages, the

[40]For this reason the rate is the paper's employment outcome outside this race. The level is reported here for the contrast.

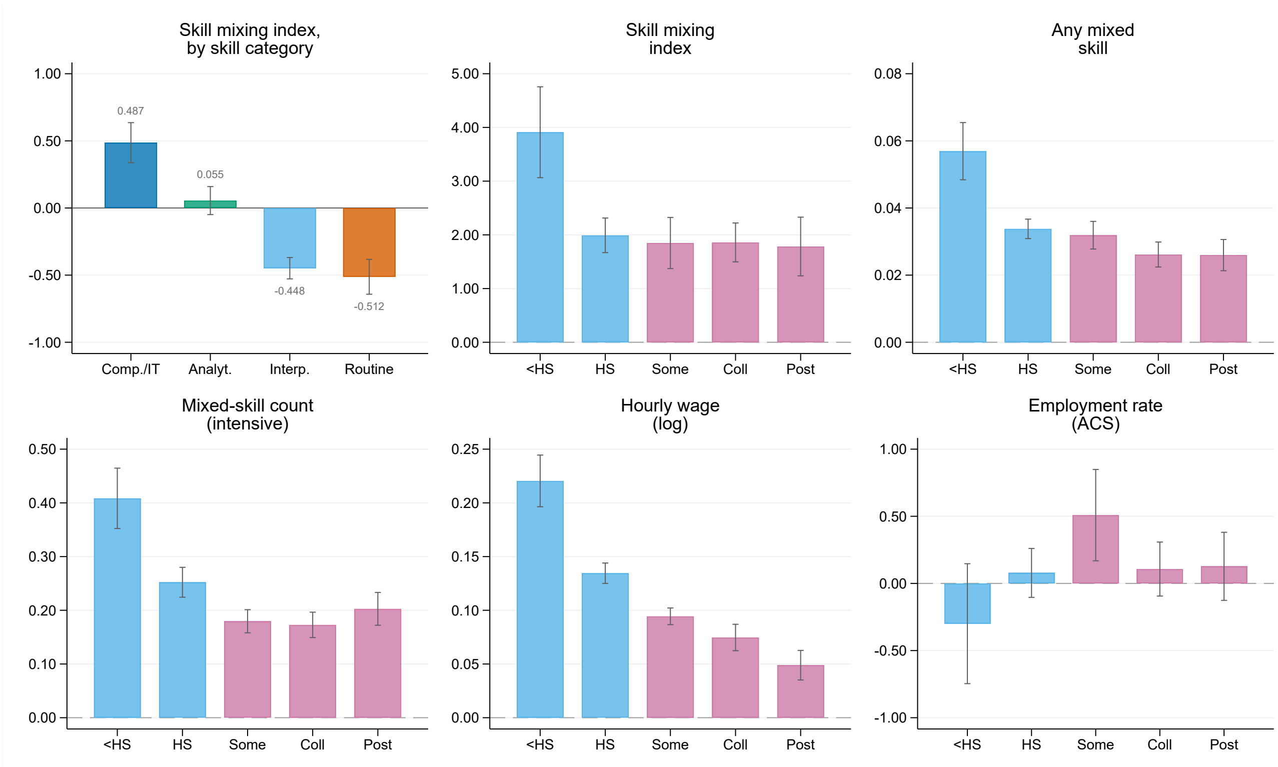


Figure 9: Heterogeneity by Skill Content and Education

Notes: This figure plots California-recentered 2SLS estimates on a 2×3 grid. The top-left panel reports the skill mixing index by skill category. For each category the outcome is the two-way cosine balance of that category against all other skills, and each bar is a separate headline 2SLS on BKGPT exposure. The remaining five panels estimate the headline specification within worker education bins, one panel per outcome, with the skill mixing index first. The employment panel is the ACS employment rate. Whiskers are 95 percent confidence intervals from standard errors two-way clustered by commuting zone and state.

any-mixed probability, and the employment level, suggesting displacement effects.

III.E Heterogeneity

I examine two margins of heterogeneity. The first margin is the skill content of the shock, and the second margin is its incidence across education levels. The first panel of Figure 9 asks which skills are being mixed in. For each skill category — computer, analytical, interpersonal, routine — the outcome is that category's balance against the rest of the bundle, the same mixing-index construction applied one category at a time. The remaining panels re-estimate the baseline 2SLS within worker-education bins. One caution on reading the first panel: these are two-way balances (each category against the rest), not the four-way headline index, so their levels are not comparable to Table 3; the panel speaks to the ordering across categories.

Skill content. Which skills does exposure mix in? The first panel re-estimates the baseline 2SLS taking each category's balance against the rest of the bundle as the outcome. A one-standard-deviation increase in exposure raises the computer balance by 0.49 and leaves the analytical balance essentially flat (0.06). The routine balance falls by 0.51. The interpersonal balance also falls.[41] This pattern is consistent with the RNR composition in Section I, where computer skills are the only group over-represented among mixed skills and routine skills are under-represented in every occupation group. Breakthrough GPTs pull computing into jobs and displace routine content in the mix.

Education gradient. The remaining panels re-estimate the baseline 2SLS within worker-education bins and show that the gains concentrate at the bottom of the education distribution. This pattern is consistent with Section I, where lower-skill occupations experience the largest secular rise in mixing. A one-standard-deviation increase in exposure raises the mixing index by 3.9 points for workers without a high school diploma, compared to 1.9 points for college and postgraduate workers. The wage effect drops from 22.0 to 4.9 log points across the same bins. The employment rate remains flat in every bin except some college, where it rises. Therefore, higher wages for less-educated workers driven by breakthrough-GPT exposure do not generate adverse employment effects.

III.F Discussion

Implications for mechanisms. The combined response of skill mixing, wages, and employment clarifies the mechanism outlined in Section IV. Exposure to Breakthrough-GPT increases both skill mixing and wages, while employment remains stable. This outcome contrasts with displacement-driven automation, which typically reduces labor demand and may decrease both employment and wages (Acemoglu and Restrepo, 2020, 2022). Instead, the findings suggest a technology-driven reorganization that enhances the value of complementary skill sets. This interpretation aligns with broader research on technology and organizational change. The adoption of GPT necessitates complementary investments in business processes, product development, and human capital (Brynjolfsson, Rock, and Syverson, 2021). Furthermore, the productivity gains from information technology are contingent upon complementary management practices (Bloom, Sadun, and Reenen, 2012).

[41] This is not the long-run rise in social-skill intensity and wage returns in Deming (2017). The panel here tracks whether breakthrough-GPT exposure tilts the mix toward interpersonal content relative to the rest of the bundle over 2010–2024 using Lightcast skills.

Limitations. Three caveats deserve mention. First, Lightcast postings capture *vacancy* skill requirements, not the tasks workers actually perform; to the extent that posting text differ from actual job task content, the mixing estimates should be read as shifts in demanded rather than realized skill bundles. Second, the patent-to-occupation mapping relies on embedding similarity and is subject to measurement error; the $\tau_0 = 0.45$ cutoff and the generality weight were fixed before any outcome regression was estimated. Appendix A.6 rebuilds exposure under alternative cosine cutoffs, each standardized by its own dispersion: the skill-mixing and wage effects remain positive and significant. Third, the design is partial-equilibrium: it identifies the effect of differential exposure. SectionV takes the mixing asymmetries as moments and asks which deep parameters can rationalize them.

IV Interpreting the Evidence

To interpret the increase in skill mixing related to breakthrough GPTs, this section develops a multidimensional matching model with endogenous skill demand. Firms design occupational skill requirements subject to the operating cost in Acemoglu (1999). The model organizes how complementarity, operating costs, and worker skill supply can produce the empirical response. It is not needed for the reduced form identification.

IV.A A Multidimensional Skill Framework

Time is discrete. At each period, there is a unit measure of heterogeneous workers who live forever. Each worker of type i is characterized by a vector of multi-dimensional skills $\mathbf{x}^i = \{x^i_1, \ldots, x^i_k, \ldots, x^i_K\} \in S \subset \mathbb{R}^{K+}$, where K is the dimension of a closed skill space S. Workers are risk-neutral and have linear utilities over consumption.

On the other side of the market, there is a mass of risk-neutral firms each running one vacancy. Firms operate different occupations $j = \{1, ..., J\}$, with $J \geq 2$. Each occupation is characterized in the same multi-dimensional skill space as workers' skills, $\mathbf{y}^j = \{y^j_1, \ldots, y^j_k, \ldots, y^j_K\} \in S \subset \mathbb{R}^{K+}$, which has the interpretation of a vector of skill requirements or skill importance for each of the worker skills.

Match output is increasing in the product of worker skill and occupational requirement in each dimension. The production function of each worker-firm pair takes a CES form of the skill inputs of workers and skill requirements of an occupation that the firm operates:

$$f(\mathbf{x}^i, \mathbf{y}^j) = \left[\sum_{k=1}^{K} (x_k^i y_k^j)^{\sigma^j} \right]^{\frac{1}{\sigma^j}}, \tag{6}$$

where σ^j is the CES across-skill substitution parameter for occupation j (the Hicks elasticity of substitution is $\varepsilon^{\text{sub}} = 1/(1 - \sigma^j)$ when $\sigma^j < 1$).[42] The parameter σ^j governs the degree of complementarity across skill dimensions. As σ^j falls, a shortfall in any one skill reduces match output more sharply, so balanced skills become more valuable. This production technology extends prior multi-dimensional skill matching models (Lise and Postel-Vinay 2020; Lindenlaub 2017) by incorporating across-skill complementarity controlled by σ^j, in addition to the multiplicative combination of worker and firm attributes. The distinction matters because breakthrough GPTs can raise skill mixing by changing the interdependence of tasks (σ), not only by shifting relative input intensities.

Further, firms actively *design* jobs subject to a cost (Acemoglu, 1999), resulting in endogenous skill demand and varying degrees of skill mixing, which offers a structural lens on its empirical rise. Firms set the occupational skill requirements $\mathbf{y}^j$, altering the production technologies that reflect the quality and optimal skill demands of the occupation. However, design is not free, and it incurs a cost $C(\mathbf{y}^j) = \sum_{k=1}^{K} (y_k^j)^{\phi^j}$ that is strictly increasing and, for $\phi^j > 1$, strictly convex in each requirement.[43] This cost captures non-wage expenses of operating an occupation that rise with skill level and job complexity.[44]

Skill mixing in the model. The firm's optimal choice of $\mathbf{y}$ determines the equilibrium degree of skill mixing, captured by the model counterpart of the skill mixing index evaluated at the optimal design $\mathbf{y}^*$. The first-order condition for job design ties each skill requirement to the worker's endowment of that skill, $y_k^* \propto x_k^c$, with the same exponent c across skills: c is the elasticity of the relative requirement to the relative skill, $d \ln(y_k^*/y_h^*)/d \ln(x_k/x_h) = c$.

[42] Throughout, "σ falls" means the substitution parameter falls, equivalently complementarity rises; I do not call σ itself an elasticity. Since labor is the only input in the model, it can be understood as "equipped" labor, and occupations' skill requirement or importance $\mathbf{y}^j$ takes a factor-augmenting form, essentially acting as demand shifters.

[43] I maintain $\phi^j > 1$ throughout, which makes the cost strictly convex ($\frac{\partial C(\mathbf{y}^j)}{\partial y_k^j} > 0$, $\frac{\partial^2 C(\mathbf{y}^j)}{\partial (y_k^j)^2} > 0$, $\forall k$). I impose the restriction because without it the design problem need not have an optimum. Profit then rises without bound along a ray in $\mathbf{y}^j$, and the first-order condition does not reach a maximum (Online Appendix B.1). The SMM estimates satisfy the restriction in both periods and occupations.

[44] For example, to operate an occupation that employs high-skill workers, a firm will need to incur higher expenses in terms of better offices and equipment rentals. Structurally, any variation in employment distribution and skill demand not explained by wages is explained by this cost.

Substituting into the mixing index therefore leaves $\mathbf{x}$ only through that common power:

$$Mix(\mathbf{y}^*) = \frac{\sum_k x_k^c}{\sqrt{K}\sqrt{\sum_k x_k^{2c}}}, \tag{7}$$

where $c = \sigma/(\phi - \sigma)$ encapsulates the interplay between technology and costs: a fall in σ or a rise in ϕ lowers c, flattening the optimal requirement vector toward equal weights. Online Appendix B.1 derives the index and its comparative statics.

The three channels. Equation (7) and Proposition 2 organize three comparative-static results. First, a fall in σ—stronger across-skill complementarity—increases the return to a weak-link skill and raises equilibrium mixing. Second, a rise in ϕ—more convex operating costs—also raises mixing because concentrating the design on a few skills becomes more expensive at the margin. Third, at fixed (σ, ϕ), evening out a single worker's skills across the K dimensions raises the mixing of the job designed for that worker—provided c lies inside a window that the estimates satisfy. What a change in the *distribution* of skills across workers does to aggregate mixing is a different question, and one the model settles by computation rather than by sign.

The first two are production parameters that a technology shock can alter. The third represents the labor supply distribution $G(\mathbf{x})$. The reduced-form rise in skill mixing aligns with expectations if breakthrough GPTs shifted a combination of these forces. Section V evaluates these mechanisms quantitatively.

V Quantification

I take the multidimensional skill matching model of Section IV to two periods of U.S. data (P_1 = 2010–12 and P_2 = 2021–23) and ask which parameters account for the responses of Section III. The quantitative skill space is three-dimensional. It uses the same Routine/Non-Routine categories of Sections I–III, with analytical and computer skills averaged into a single analytical/computer dimension alongside interpersonal and routine skills ($K = 3$; Appendix B.3).[45] Two estimated forces govern equilibrium mixing in this space. The first is skill substitutability within an occupation's task bundle, governed by the CES parameter σ (where a lower σ implies greater complementarity). The second is the convexity of the cost of operating a mixed-skill design, governed by ϕ.[46] These operate alongside the third channel of

[45]The merge keeps the quantitative problem tractable while preserving the four-category content facts of the reduced form; throughout I refer to the combined dimension as "analytical/computer."

[46]Terminology follows Section IV. "σ falls" or "σ-drop" refers to a decline in the substitution parameter, equivalently an increase in skill complementarity; I avoid the phrase "elasticity of substitution" for σ itself.

Table 5: Mapping education estimates into calibration targets

	Slope per SD of exposure		2010 ACS weight	
Education group	Mixing	Wages	w_e^H	w_e^L
Less than high school	3.91	22.05	0.0177	0.1317
High school graduate	1.99	13.45	0.1322	0.3939
Some college	1.85	9.44	0.2697	0.3552
College graduate	1.86	7.47	0.3291	0.1009
Postgraduate	1.78	4.89	0.2513	0.0183
	Mixing in index points		Wages in log points	
	High wage	Low wage	High wage	Low wage
Weighted response per SD	1.891	2.176	8.401	12.398
Response per era dose	**5.33**†	**6.13**†	23.68	34.95
Targeted wage gap response			$-$**11.27**†	

Notes: This table maps the five education-bin Bartik estimates of Figure 9 into the model's low- and high-wage occupation targets. Each occupation response is a composition-weighted average of those estimates, using fixed 2010 ACS education shares within the occupation group. The upper panel reports the education-bin slopes per one standard deviation of exposure and the corresponding weights. The lower panel reports the weighted averages per standard deviation and the same objects scaled by the era dose of 53.7 patent units (2.819 cross-market standard deviations). The wage-gap target is the high-wage minus low-wage wage response. Daggers mark the three Section III objects that enter the calibration.

Section IV, the externally calibrated worker skill supply $G(\mathbf{x})$.

V.A Calibration

External calibration. Matching frictions, discounting, and skill-adjustment speeds are held at standard values (Appendix Table B1, Panels A–B).[47] The worker share of match output is $\omega = 0.6$, matching the U.S. labor share. The worker skill supply $G(\mathbf{x})$ is calibrated from NLSY period by period (Appendix B.4) and held outside the estimation. The technology the data are asked to move is described entirely by (σ, ϕ).

Internal estimation. I estimate the eight production parameters $((\sigma_1^{P_t}, \sigma_2^{P_t}, \phi_1^{P_t}, \phi_2^{P_t})$ for $t = 1, 2)$ by simulated method of moments, using five targets in each period. Three come from the estimates in Section III. Table 5 recasts the five education-bin responses for the model's low- and high-wage occupations, using their 2010 ACS employment composition, and scales them by the average market's exposure. They are the two mixing responses and the wage-gap response. Exposure raises mixing in each occupation and, because wages rise by different amounts, changes their wage gap. The two aggregate moments are the high-wage employment share and the unemployment rate (Appendix Table B2).

[47] Sources as in the table notes (Shimer (2005), Mercan and Schoefer (2020), Braxton, Herkenhoff, and Phillips (2020), and Lise and Postel-Vinay (2020)).

Table 6: Calibrated production parameters

Parameter	Value
First environment	
Skill substitution in low wage occupations $\sigma_1^{P_1}$	0.898
Skill substitution in high wage occupations $\sigma_2^{P_1}$	0.751
Cost convexity in low wage occupations $\phi_1^{P_1}$	3.587
Cost convexity in high wage occupations $\phi_2^{P_1}$	2.942
Second environment	
Skill substitution in low wage occupations $\sigma_1^{P_2}$	0.815
Skill substitution in high wage occupations $\sigma_2^{P_2}$	0.737
Cost convexity in low wage occupations $\phi_1^{P_2}$	3.752
Cost convexity in high wage occupations $\phi_2^{P_2}$	3.476
Change in the substitution parameter	
Low wage occupations $\Delta\sigma_1$	0.083
High wage occupations $\Delta\sigma_2$	0.014

Notes: This table reports the response calibration of the production parameters. The criterion uses the two projected mixing responses, the occupational wage gap response, the high wage employment share, and unemployment in each environment. The two environment counterparts of each exposure response share one pooled empirical estimate. The parameter vector is therefore interpreted as a calibration rather than as a conventional overidentified estimator. No sampling intervals are reported because the discrete bundle grid produces unstable finite difference derivatives.

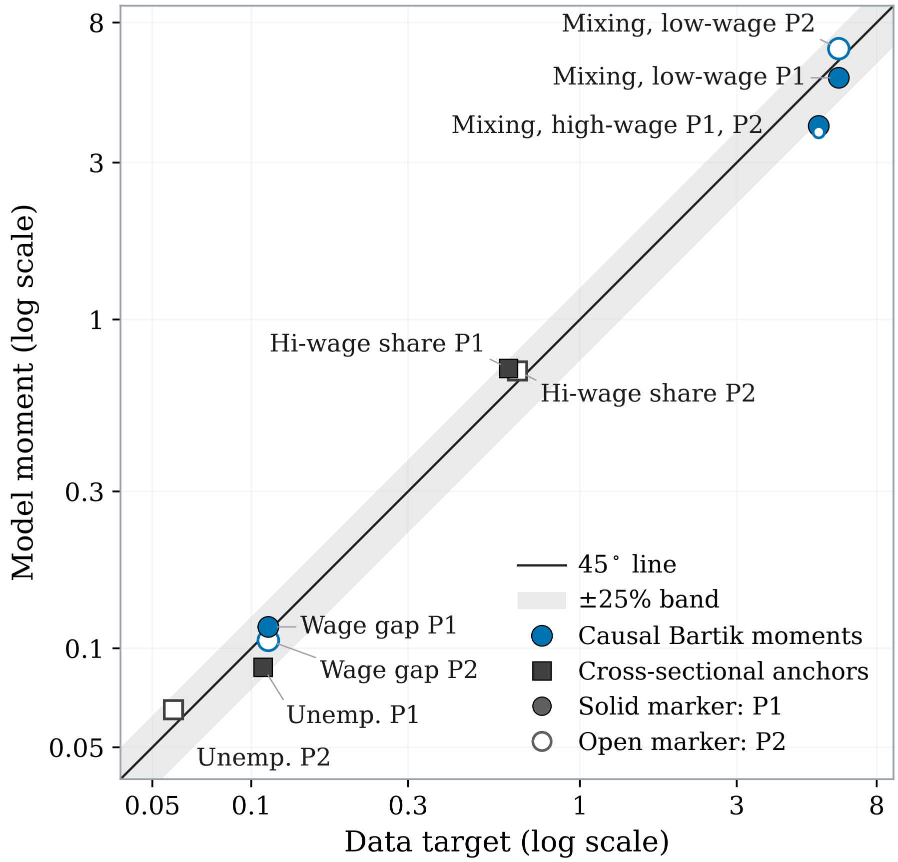
Identification. The mixing responses identify the substitution parameters. When complementarity deepens, an occupation whose bundle starts narrow must rebalance, and the size of each occupation's response pins its σ-drop. The wage gap identifies the cost side jointly with the anchors. Given the mixing response, how much of the redesign's surplus reaches the low-wage occupation depends on the cost of operating the redesigned bundle, so the gap response disciplines (ϕ_1, ϕ_2). It can do so where the wage levels cannot, because the levels move with the efficiency normalization while the gap is free of it.

V.B Estimated Technology Shifts

Which production parameters moved between P_1 and P_2, and by how much? Table 6 reports the eight estimates and the implied complementarity deepenings $\Delta\sigma_o = \sigma_o^{P_1} - \sigma_o^{P_2}$; Figure 10 places every targeted moment against its data counterpart.

Table 6 shows that skill complementarity deepens in both occupations, but more in the low-wage occupation. σ_1 falls from 0.898 to 0.815 ($\Delta\sigma_1 = 0.083$), whereas σ_2 falls from 0.751 to 0.737 ($\Delta\sigma_2 = 0.014$). In elasticity units, the low-wage occupation's skill inputs become about half as substitutable ($\varepsilon = 1/(1-\sigma)$ falls 44.8 percent, from 9.77 to 5.39), compared with a

Figure 10: Calibration fit



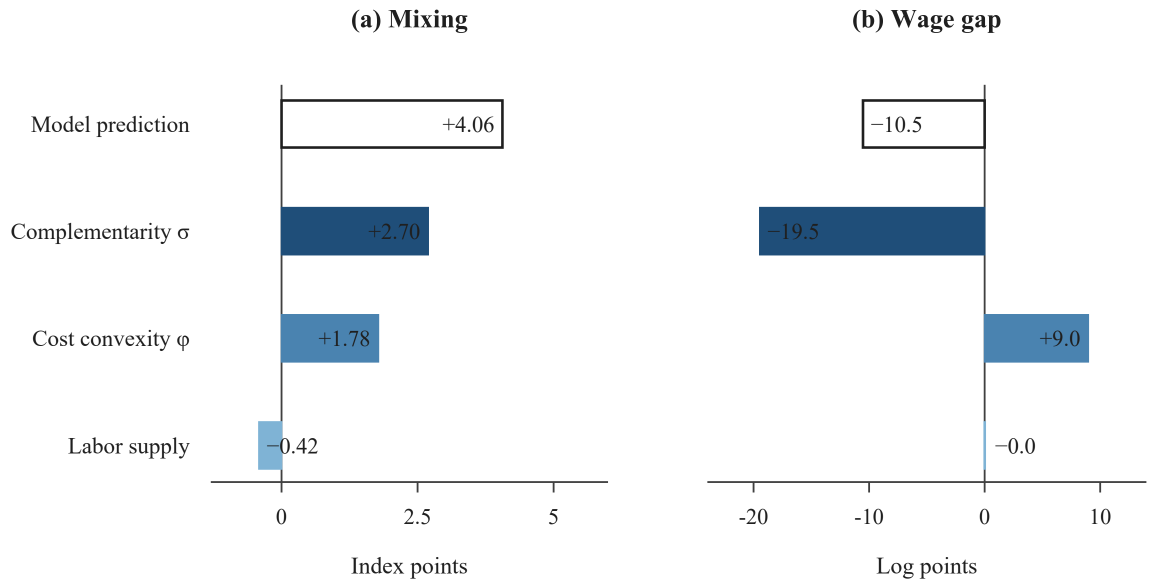
Notes: This figure plots each model condition against its calibration target on logarithmic axes. The diagonal is exact agreement and the shaded region lies within twenty five percent of it. Blue circles denote the exposure-response conditions and grey squares denote the cross sectional anchors. Filled markers refer to the first calibrated environment and open markers to the second.

five percent decline in the high-wage occupation (4.01 to 3.80). As σ falls, a weak link in any one skill drags down the whole product, so a mixed skill set becomes more valuable. The low-wage deepening therefore makes a narrow design more costly where skills were initially more substitutable.

The table also shows that the cost of operating mixed-skill designs rises in both occupations, and more steeply in the high-wage one (ϕ_1 from 3.587 to 3.752; ϕ_2 from 2.942 to 3.476). A higher ϕ itself pushes task bundles toward balance (Proposition 2), so complementarity and operating costs can move mixing in the same direction. The next subsection asks how much each channel contributes.

Fit. Figure 10 compares each target with its model counterpart. Eight of the ten moments lie within 25 percent of their targets. The model closely matches the wage-gap response and the low-wage mixing response. Its main miss is high-wage mixing, which is about three-tenths

Figure 11: Accounting for mixing and the occupational wage gap



Notes: This figure plots an accounting decomposition of the predicted net change. The open bar is the net change predicted by the model. The filled bars report the signed contributions of skill complementarity, cost convexity, and labor supply. The last component is worker skill supply $G(\mathbf{x})$; the bar also carries the calibrated posting-cost path. Every state behind a bar is a solved equilibrium and the three components sum to the open bar apart from rounding. Panel A reports mixing in index points. Panel B reports the occupational wage gap in log points. The calculation is conditional on the calibrated parameter vector. It is an accounting decomposition and does not separately identify a causal contribution for either production channel.

below its target in both environments. The fit therefore supports the low-wage-led ordering of mixing magnitudes. Appendix Table B2 reports every moment.

V.C Decomposing Skill Mixing and Labor Market Outcomes

What accounts for the estimated movements in skill mixing between the two periods, and how do the same forces shape the occupational wage gap? Figure 11 decomposes the model's $P_1 \to P_2$ response at the era dose into three channels—skill complementarity σ, cost convexity ϕ, and labor supply $G(\mathbf{x})$. Panel (a) reports the contributions to mixing, and panel (b) the contributions to the occupational wage gap.

Mixing. I begin with the aggregate mixing response in Panel (a). The estimated shift raises skill mixing by 4.1 index points. The subsequent decomposition shows that the deepening of skill complementarity and the rise in operating costs account for the increase. Complementarity contributes 2.7 points and the cost shift adds 1.8, so complementarity carries most of the net move. Changes in worker skill supply do not drive this rise; these shifts subtract 0.4 points. These results are consistent with the predictions of Section IV and highlight the role of complementarity and operating costs in firms' endogenous redesign of skill bundles.

Wage-gap fit. I proceed to investigate how the same channels that drive skill mixing also shape the occupational wage gap. Section III shows that, at the dose carried by the average market, exposure raises wages by 34.9 log points in low-wage occupations and 23.7 in high-wage occupations, compressing their gap by 11.3 log points. Panel (b) of Figure 11 reports the corresponding decomposition. The growing complementarity of skills accounts for the decline in the gap. On its own it would compress the gap by roughly three-quarters again as much. Higher cost convexity works in the opposite direction and gives back nearly half of that compression, as more convex operating costs make it costlier to run those redesigned jobs. As with skill mixing, changes in worker skill supply and posting costs do not drive the result. In short, deeper complementarity therefore dominates both decompositions. It accounts for most of the mixing response, while cost convexity offsets part of the wage-gap compression.

VI Concluding Remarks

Given the multidimensional nature of modern work, understanding employers' demand for skill *mixtures* is crucial for explaining labor market dynamics. Leveraging a near-universe of online job postings and longitudinal task data, I document a structural shift toward skill mixing—the hierarchical broadening of job roles in which high-skill, non-routine capabilities are increasingly integrated into lower-skill occupations. Nearly half of the skills an occupation lists in 2024 were absent from it in the early 2010s, and almost all of those arrivals were already demanded elsewhere. Such mixing is also reflected in the rise in the skill mixing index. Using patent and vacancy data from 2010 to 2024, I establish breakthrough GPTs as a key driver of this phenomenon. Exposure to these technologies raises skill mixing and wages and leaves the employment rate essentially unchanged, with gains concentrated among less-educated workers. Through the lens of a multi-dimensional matching model with endogenous occupation design, increased skill complementarity together with more convex occupational operating costs accounts for the rise in mixing and the wage response.

These findings carry implications for human capital policy. While this paper focuses on the demand side, future work could explore the supply side—specifically, how educational curricula and training programs can be redesigned to foster multidimensional skill sets, so that workers are equipped to reap the returns from the changing demand for skill mixtures.

References

Acemoglu, D. (1999). Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review*, 89(5):1259–1278.

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Acemoglu, D. and Restrepo, P. (2022a). Demographics and automation. *The Review of Economic Studies*, 89(1):1–44.

Acemoglu, D. and Restrepo, P. (2022b). Tasks, automation, and the rise in US wage inequality. *Econometrica*, 90(5):1973–2016.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *Quarterly Journal of Economics*, 134(4):1949–2010.

Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The evolution of work in the united states. *American Economic Journal: Applied Economics*, 12(2):1–34.

Atalay, E. and Sarada (2020). Emerging and disappearing work, thriving and declining firms. *Working Paper*.

Autor, D., Chin, C., Salomons, A., and Seegmiller, B. (2024). New frontiers: The origins and content of new work, 1940–2018. *The Quarterly Journal of Economics*, 139(3):1399–1465.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–97.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China syndrome: local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168.

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

Autor, D. H. and Price, B. (2013). The changing task composition of the us labor market: An update of autor, levy, and murnane (2003). *Working Paper*.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? *WE Upjohn Institute for Employment Research*.

Bloom, N., Sadun, R., and Reenen, J. V. (2012). Americans do it better: Us multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Borusyak, K. and Hull, P. (2023). Non-random exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213.

Braxton, J. C., Herkenhoff, K. F., and Phillips, G. M. (2020). Can the unemployed borrow? implications for public insurance. Technical report, National Bureau of Economic Research.

Braxton, J. C. and Taska, B. (2023). Technological change and the consequences of job loss. *American Economic Review*, 113(2):279–316.

Bresnahan, T., Greenstein, S., Brownstone, D., and Flamm, K. (1996). Technical progress and co-invention in computing and in the uses of computers. *Brookings Papers on Economic Activity. Microeconomics*, 1996:1–83.

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

Brynjolfsson, E., Rock, D., and Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1):333–372.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *Quarterly Journal of Economics*, 139(2):891–936.

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

Crafts, N. (2004). Steam as a general purpose technology: a growth accounting perspective. *The Economic Journal*, 114(495):338–351.

De Chaisemartin, C. and d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, pages 1–45.

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(4):1593–1640.

Freeman, R. B., Ganguli, I., and Handel, M. J. (2020). Within-occupation changes dominate changes in what workers do: A shift-share decomposition, 2005–2015. In *Aea Papers and Proceedings*, volume 110, pages 394–99.

Goldin, C. and Katz, L. F. (2010). *The race between education and technology*. harvard university press.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10(1):92–116.

Hall, B. H., Jaffe, A., and Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 36(1):16–38.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools.

Helpman, E. and Trajtenberg, M. (1994). A time to sow and a time to reap: Growth based on general purpose technologies.

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

Institute for Structural Research (2016). Occupation classifications crosswalks: From O*NET-SOC to ISCO. Online database and Stata code, https://ibs.org.pl/en/resources/occupation-classifications-crosswalks-from-onet-soc-to-isco/. Accessed August 13, 2026.

Kalyani, A., Bloom, N., Carvalho, M., Hassan, T., Lerner, J., and Tahoun, A. (2025). The diffusion of new technologies. *The Quarterly Journal of Economics*, 140(2):1299–1365.

Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring technological innovation over the long run. *American Economic Review: Insights*, 3(3):303–320.

Kim, G., Merritt, C., and Peri, G. (2024). Measuring and predicting "new work" in the united states: The role of local factors and global shocks. Technical report, National Bureau of Economic Research.

Lightcast (2024). Representativeness analysis of lightcast job posting data - u.s. Technical report, Lightcast.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, 84(2):718–789.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–76.

Mercan, Y. and Schoefer, B. (2020). Jobs and matches: Quits, replacement hiring, and vacancy chains and vacancy chains. *American Economic Review: Insights*, 2(1):101–124.

Ocampo, S. (2022). A task-based theory of occupations with multidimensional heterogeneity. *Working Paper*.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, 11(1):99–126.

Tsvetkova, A., D'Amico, E., Lembcke, A., Knutsson, P., and Vermeulen, W. (2024). How well do online job postings match national sources in large english speaking countries? benchmarking lightcast data against statistical sources across regions, sectors and occupations. *OECD Local Economic and Employment Development (LEED) Working Papers*, (1):I–65.

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

Appendix for Online Publication

Table of Contents

A ADDITIONAL EMPIRICAL RESULTS

A.1 Data Construction

Lightcast provides the job vacancy text analyzed in Sections I and III, while O*NET functions as an independent validation source for the intensive margin.

Lightcast gathers daily online job vacancies from more than 51,000 sources, such as job boards, employer websites, public portals, and staffing agencies. It collects information on employers, occupations, locations, wages, and skills, and identifies both new and active job ads (Lightcast, 2024). The dataset covers 2010 to 2024. To avoid counting duplicates, repeated ads within 60 days are combined. Skills come from the public Open Skills library. About 2,000 similar skill labels were merged using transformer similarity, but the original skill identifiers were kept to maintain links to the source taxonomy (Kim, Merritt, and Peri, 2024).

Lightcast data matches official sources closely. Since 2013, active postings have averaged 92.6 percent of JOLTS openings, with a time series correlation of 0.87. State shares have a 0.98 correlation with OEWS employment shares, and broad occupation shares correlate at 0.74 (Lightcast, 2024). The main limitation is that small firms and offline hires are underrepresented, which affects overall levels but probably does not bias changes within cells. This analysis uses 2010 ACS employment shares, checks trends with raw text, and looks at frequency, vocabulary, employer, and posting volume in Appendix A.4.

O*NET provides an independent check on the intensive margin because its data are derived from worker surveys regarding job context, activities, and knowledge, rather than job postings. Each release updates only a portion of the approximately 970 detailed occupations. Survey waves in which 54 to 74 percent of occupations were updated after the reference year are used, treating the series as indicative of gradual changes in occupation content rather than as an annual panel. SOC codes are matched to Census occupations using official crosswalks and the Autor Dorn panel updated by Deming (2017). Table A1 presents cumulative update coverage for each wave.

Table A1: O*NET Versions and Corresponding Years

Version	Released Year	Division	Work Context	Work Activities	Knowledge	Skills	Abilities	Considered Year
O*NET 13.0	2008	Post 2005	73.79%	73.79%	73.79%	73.79%	73.79%	2005
		Before 2005	26.21%	26.21%	26.21%	26.21%	26.21%	
O*NET 18.0	2013	Post 2009	57.15%	57.21%	57.21%	99.89%	57.21%	2009
		Before 2009	42.85%	42.79%	42.79%	0.11%	42.79%	
O*NET 22.0	2017	Post 2013	57.84%	57.67%	57.67%	57.67%	57.67%	2013
		Before 2013	42.16%	42.33%	42.33%	42.33%	42.33%	
O*NET 25.0	2022	Post 2018	54.52%	54.52%	54.52%	54.52%	54.52%	2018
		Before 2018	45.48%	45.48%	45.48%	45.48%	45.48%	

Notes: This table reports O*NET database versions, their release year, the year division for the five modules (work context, work activities, knowledge, skills, abilities), and the considered year for each version. The Post and Before percentages are cumulative shares of seven-digit occupations whose survey data were collected after or before the indicated cutoff within each version. They are not the per-release refresh rate, which is about 110 of 970 occupations per rolling update. The Considered Year is chosen so that a majority of occupations have post-cutoff data in the modules used for the skill-mixing cross-check.

Table A2: Lightcast RNR keyword lexicon and assignment routes

RNR category	Representative keywords	Source
Analytical	*research*, *analy*, *decision*, *solving*, *math*, *statistic*, *thinking*	Braxton and Taska (2023); Hershbein and Kahn (2018)
Interpersonal	*communication*, *teamwork*, *collaboration*, *negotiation*, *presentation*	Braxton and Taska (2023); Deming–Kahn (2018)
Computer	*computer* on parsed skill tags, or any Lightcast software skill; body-text route: software list only	Lightcast taxonomy
Routine cognitive	*bookkeep*, *calculate*, *calculating*, *correct*, *correcting*, *corrections*, *measurement*, *measuring*	Hershbein and Kahn (2018)
Routine manual	*control*, *controlling*, *equip*, *equipment*, *equipping*, *operate*, *operating*	Hershbein and Kahn (2018)
Routine (index)	Mean of routine-cognitive and routine-manual flags	—

Notes: This table reports the keyword lexicon used on both Lightcast routes and matches the replication code (`skill_mix.py`; `3.2 measure_skill_mix.do`). Named-skill route (Section I, the skill mixing index): software-flagged skills map to Computer; the keyword lists cover the remaining categories; what neither layer names is assigned to the nearest category centroid in embedding space, only above a calibrated similarity threshold, so skills clearing no layer are left uncategorized. Posting-level keyword route (Figure 4 panel (b) gray series and Figure 5 panel (a)): the same keywords are case-insensitive substring matches on skill-tag strings or vacancy body text. O*NET work-context descriptors in Table A4 are used only for the O*NET cross-check of Section I, not for Lightcast RNR tagging.

A.2 Details of Skill Measures

The analysis separates the occurrence and extent of skill recombination. Count measures treat each consolidated Lightcast skill as its own category and note when a requirement appears outside its original occupation. The four-skill index checks if the skill bundle is balanced across analytical, interpersonal, computer, and routine content. Lightcast software flags mark computer skills, while the Hershbein Kahn lexicon classifies other named skills. An embedding layer assigns a skill only if its similarity meets a threshold set for 0.90 agreement in held-out classification. Skills below this threshold stay unclassified; they are counted in totals but left out of the four-skill index. A third method searches the raw vacancy text directly. Across these methods, the eight annual series have correlations above 0.62, and mostly above 0.85.

Table A3 shows how posting-level detection margins relate to the skill mixing index. At the commuting-zone by demographic-cell level, the mixed-skill count and the index have a 0.86 correlation in levels and 0.71 in within-cell changes from 2011 to 2023. Both measures

Table A3: Coherence of the skill-mixing measures at the cell level

Measure pair	Correlation in levels	Correlation in changes
Mixed-skill count, skill mixing index	0.86	0.71

Notes: This table reports employment-weighted correlations at the commuting-zone $\times$ demographic-cell level. Levels pool all cell-years. Changes are within-cell long differences from 2011 to 2023. The pair is the mixed-skill count and the skill mixing index of Table 3.

show the same trend: the count shows when mixing happens, while the index measures the shift in skill demand. In the cumulative-exposure design, the estimated responses for the count and the index are similar (Table 3).

O*NET offers a secondary check on the mixing patterns seen in Lightcast data. Table A4 lists the O*NET work-context and work-activity descriptors used for this. The analytical and interpersonal measures match the "non-routine cognitive analytic" and "non-routine interpersonal" tasks from Acemoglu and Autor (2011). I combine their "routine cognitive" and "routine manual" tasks into one routine skill, since jobs needing either show similar labor-market patterns (Autor, Levy, and Murnane, 2003; Acemoglu and Autor, 2011). "Non-routine manual" is left out because its descriptors come from the "Abilities" module, which is rated by analysts, not workers.

Table A4: O*NET Skill Measures and Composing Descriptors

Skill Category	Task Descriptors
Non-routine analytical	Analyzing data/information
	Thinking creatively
	Interpreting information for others
Non-routine interpersonal	Establishing and maintaining personal relationships
	Guiding, directing and motivating subordinates
	Coaching/developing others
Computer	Interacting with computers
	Programming
	Computers and electronics
Routine	Importance of repeating the same tasks
	Importance of being exact or accurate
	Structured vs. unstructured work (reverse)
	Pace determined by speed of equipment
	Controlling machines and processes
	Spend time making repetitive motions

Notes: This table lists the O*NET descriptors behind the skill measures. Non-routine Analytical and Non-routine Interpersonal align with Acemoglu and Autor (2011)'s non-routine cognitive analytic and non-routine interpersonal skills. A unified Routine measure combines that paper's routine cognitive and routine manual skills. Non-routine manual is omitted so that the set stays with incumbent worker-evaluated descriptors.

A.3 Robustness of Skill Mixing Trends

This subsection presents two robustness checks for the main skill-mixing trends in Section I. One uses a different occupational classification, and the other shows the annual paths of the main-text decomposition totals in percentile-rank units.

Figure A1 replicates the four-category skill breakdown from Section I.A, utilizing the Autor, Dorn, and Deming harmonized Census occupation panel in place of the 4-digit ISCO codes from Figure 1. As the two classifications are independent, this serves as a robustness check employing an alternative classification rather than a more detailed one. Both panels employ Census codes: Panel (a) displays skill counts by major group, and Panel (b) presents the share of postings with at least one mixed skill using the same mapping. The observed pattern remains largely unchanged. Mixed skills continue to represent the primary entry channel, and the cross-occupation gradient persists. The posting-share comparison maintains the ordering, though with a reduced gap (35 percent on average in lower-wage major groups and 31 percent in higher-wage groups under Census codes, compared to 56 percent and 34 percent at the 4-digit ISCO level).

Figure A2 shows the annual series behind the totals in main-text Table 1. The skill mixing index and the keyword-search index are reported as employment-weighted mean percentile ranks of their 2011 cell distributions, reaching about the 59th and 56th percentiles by 2024. The O*NET survey index is shown as the employment-weighted mean rank of the 2005 distribution at each wave, reaching about the 60th percentile by 2018.

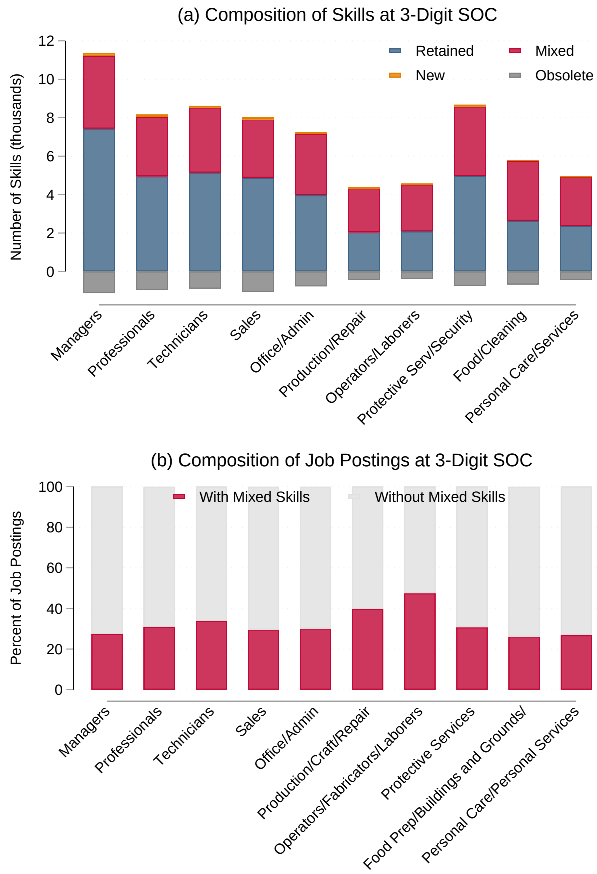
A.4 Lightcast Taxonomy Stability and Raw-Text Reconstruction

The baseline analysis uses a dynamic set of vacancies and a curated skill taxonomy. This subsection assesses whether increased measurement from either source could explain the observed mixing trend, even if employers' stated requirements remain unchanged.

Eight series use various combinations of parsed tags, raw text, posting shares, and distinct counts. Seven of these series increase from 2011 to 2024. Pairwise correlations exceed 0.62, and are mostly above 0.85. This indicates that the pattern is present in the raw text regardless of changes in the taxonomy.

Table A5 evaluates different frequency thresholds, posting share thresholds, base years, vocabularies, employers, and occupation-employer pairs. Frequency rules remove rare mentions, while posting share rules maintain consistent detection as vacancy volume increases. In both cases, mixed skills account for 41.4 to 51.9 percent of 2024 skills and 95.0 to 98.5 percent of new skills. Using 2011 or 2012 as the base year reduces the mixed share to 40.4

Figure A1: Composition of Skills in US Job Postings under Census Occupation Codes, 2010 to 2024

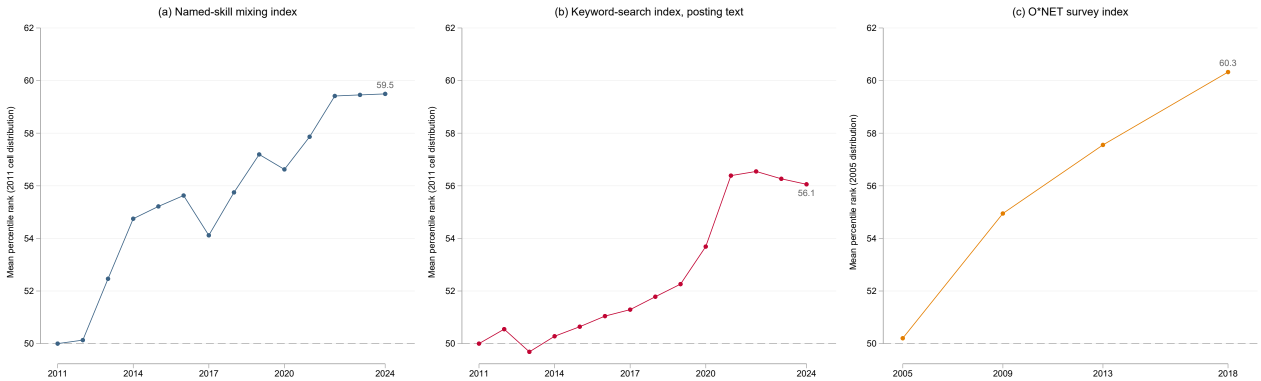


Notes: This figure plots the composition of skills under Autor–Dorn–Deming Census occupation codes. Skills are classified at the 3-digit `occ1990dd` level and averaged within major groups. Panel (a) splits skills into four categories based on emergence and persistence. Values are in thousands, and declining skills sit below the horizontal axis. Panel (b) plots the share of postings in each major Census group that list at least one mixed skill, using the same classification and aggregation.

and 38.8 percent, as more skills are already present. Freezing the vocabulary keeps the mixed share at 46.7 percent and reduces new skills from 1.9 to 0.5 percent. Balancing by employers and employer-occupation pairs results in 44.4 and 42.7 percent, respectively. These checks distinguish the spread of existing skills from increases in labels, posting volume, and employer entry.

Limiting the sample to 9,891 skills with an external software or keyword name results in a mixed share of 45.2 percent, compared to 46.1 percent in the full set. The occupation ranking remains unchanged. These checks confirm that the trend is driven by recombining existing

Figure A2: Annual Index Trends in Percentile-Rank Units



Notes: This figure restates the three mixing-index series of Figures 4(b) and 5 in the percentile-rank units of Table 1. In panels (a) and (b), each commuting-zone × four-digit ISCO-08 cell's index is converted to its percentile rank in that measure's 2011 cell distribution (tie-aware midranks over the employment-weighted cell distribution; values outside the 2011 support are ranked by the right-continuous weighted CDF). The plotted series is the employment-weighted mean rank by year, equal to 50 in 2011 by construction. Panel (c) plots the O*NET survey index at waves 2005, 2009, 2013, and 2018 as the ACS employment-weighted mean of harmonized Census occupations' percentile ranks in the 2005 distribution. End-of-series labels report the final-year mean rank. The dashed line marks the anchor-year mean of 50. The 2011–2024 and 2005–2018 changes are the Totals decomposed in Table 1.

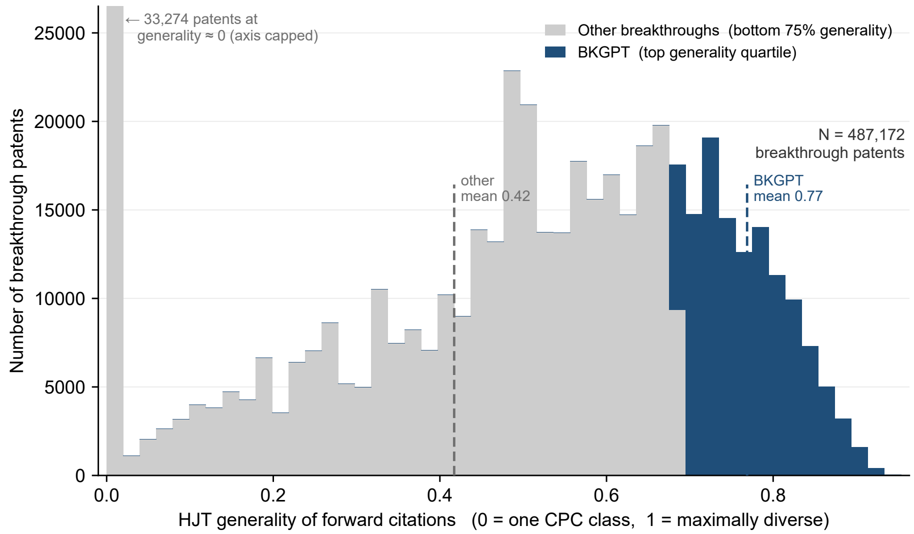
skills, not by changes in the taxonomy or relabeling.

Table A5: Skill Decomposition Under Presence Rules, Base Years, and Sample Restrictions

Presence rule	% of skills observed in 2024: New to occ.	Mixed	Genuinely new	Obsolete (%)	Mixed share of new (%)
Panel A. Presence rule (base year 2010)					
≥1 posting (baseline, Fig. 1)	48.0	46.1	1.9	18.8	96.0
≥10 postings	53.2	51.9	1.3	12.1	97.5
≥0.1% of occupation's postings	42.1	41.4	0.6	24.5	98.5
≥1 posting, occs. above median 2010 volume	44.5	42.3	2.2	17.3	95.0
Panel B. Base year of the decomposition (≥1 posting rule)					
Base year 2011	42.4	40.4	2.0	22.2	95.3
Base year 2012	40.8	38.8	2.1	23.2	95.0
Panel C. Sample restriction (base year 2010, ≥1 posting rule, 2025-06 Databricks vintage)					
Frozen 2010–11 vocabulary	47.2	46.7	0.5	19.4	99.0
Balanced employers (2010–12 ∩ 2022–24)	46.3	44.4	1.9	21.3	96.0
Balanced employer × occupation	44.5	42.7	1.8	23.1	95.9

Notes: This table reports the four-category skill decomposition of Figure 1. Panel A varies the presence rule for detecting a skill in a four-digit ISCO occupation-period, applied symmetrically in 2010 and 2024. Share rules scale the threshold by the occupation-period's posting volume, so presence is invariant to volume growth. The last row of Panel A keeps the 203 occupations (of 411) with 2010 posting volume above the median (7,284 postings). Panel B replaces the 2010 base year with 2011 or 2012 and holds the 2024 target and the ≥1 posting rule fixed. A later base year lowers the extensive count because more skills are already on record. Panel C holds the presence rule and base year at baseline and restricts the sample to the 2010–11 vocabulary, to employers in both 2010–12 and 2022–24, and to employer × two-digit occupation pairs in both windows, on the June 2025 Lightcast pull, which matches the Panel A mixed share to within 0.06 percentage points (46.05 against 46.11). *New to occ.* is the share of 2024 skills absent from the occupation in the base year; *Mixed* are those present elsewhere in the base year; *Genuinely new* are globally new; *Obsolete* is the share of base-year skills absent by 2024; *Mixed share of new* is Mixed/(Mixed + Genuinely new). Shares pool occupation-skill pairs across four-digit ISCO occupations, excluding armed forces. Figure 1 averages the same units within one-digit groups. Source: Lightcast, 2010 and 2024.

Figure A3: Generality of breakthrough patents



Notes: This figure plots the Hall–Jaffe–Trajtenberg generality index among well cited breakthrough patents. The dark distribution contains the top generality quartile used to define breakthrough GPT patents.

A.5 Breakthrough-GPT Patents: Additional Detail

Figure A3 presents the distribution of generality among well-cited breakthroughs. The top quartile identifies patents with the broadest impact.

A.6 Patent–O*NET Mapping Validation

The exposure measure is determined prior to its combination with local occupation shares. This subsection details the process by which patent text is converted into an occupation shock and evaluates whether reasonable modifications to this mapping influence the results. Patent-occupation pairs are retained when cosine similarity exceeds $\tau_0 = 0.45$. The weight is defined as $(\cos\text{-sim} - \tau_0)G_p$, where G_p denotes citation generality. Subtracting the threshold increases the influence of pairs with stronger textual correspondence while ensuring zero weight at the boundary. The retained matches are crosswalked to ISCO 2 and accumulated into the exposure measure. This continuous rule averages the relevant occupation matches, rather than assigning each patent to a fixed number of jobs.

Coverage Of 82,052 breakthrough GPT patents, 53,693 have an embedding that clears the crosswalk. Under the continuous rule at τ_0=0.45, 6,593 patents match at least one ISCO-2 occupation (12.3 percent of the cosine pipeline), contributing to 2.8 occupations on average and one at the median. Tighter thresholds reduce coverage monotonically.

Table A6: Patent-to-Occupation Matching Coverage Across Continuous Cosine Thresholds (τ_0)

	τ_0=0.40	τ_0=0.45	τ_0=0.50	τ_0=0.55
Matched BK–GPT patents (count)	18,568	6,593	1,678	365
Share of cosine pipeline (%)	34.58	12.28	3.13	0.68
Median # ISCO-2 per patent	2	1	1	1
Mean # ISCO-2 per patent	3.38	2.80	2.62	2.48
Share with ≥ 2 matches (%)	60.12	49.52	43.86	45.75
Share with > 10 matches (%)	5.88	4.46	4.65	3.84

Notes: This table reports coverage diagnostics for the paper's continuous matching rule at alternative cosine thresholds $\tau_0 \in \{0.40, 0.45, 0.50, 0.55\}$ (no top-K cap). The breakthrough-GPT patent universe is the 82,052 patents with the citation-free text-based BK flag and 2010–2024 filing date. Of these, 53,693 (65.4 percent) carry at least one patent×O*NET skill embedding with a survivable SOC–ISCO crosswalk and enter the matching pipeline; that cosine-pipeline denominator is used in the Share of cosine pipeline row. A patent is matched at a given τ_0 when at least one ISCO-2 occupation clears the cosine cutoff; occupation counts are the distinct ISCO-2 matches above τ_0 for that patent. The column τ_0=0.45 is the headline cutoff. Coefficient robustness across the same thresholds is in Table A7.

Table A7: Headline Coefficients Across Continuous Cosine Thresholds (τ_0)

	τ_0=0.40	τ_0=0.45	τ_0=0.50	τ_0=0.55
Panel 1. Skill mixing index				
	1.51***	1.89***	2.32***	2.00***
	(0.17)	(0.21)	(0.26)	(0.22)
Panel 2. Log hourly wage				
	0.123***	0.140***	0.149***	0.125***
	(0.008)	(0.010)	(0.012)	(0.010)
Panel 3. Employment rate (ACS)				
	0.020	0.026	0.015	0.062
	(0.090)	(0.093)	(0.093)	(0.078)

Notes: This table reports 2SLS coefficients on the cell-level breakthrough-GPT Bartik under the paper's continuous matching weight $(\cos\text{-sim} - \tau_0) \cdot G_p$, rebuilt at alternative cosine thresholds $\tau_0 \in \{0.40, 0.45, 0.50, 0.55\}$. Matched patent–skill intensities are aggregated to ISCO-2 and cumulated as in the headline exposure. The column τ_0=0.45 recovers Table 3, column 5. The IV is the California year-demeaned leave-out instrument. Specification, controls, fixed effects (commuting zone, demographic cell×year, state×year), weights, and clustering match the headline. Every column is per 1 SD of that column's 2010-employment-weighted exposure. Standard errors clustered by commuting zone and state. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Cosine cutoff Table A7 rebuilds the continuous weight at $\tau_0 \in \{0.40, 0.45, 0.50, 0.55\}$. Each column uses its own exposure standard deviation and the baseline regression design. Mixing and wages are positive and significant at one percent in every column. Mixing ranges from 1.51 to 2.32 and wages from 0.123 to 0.149, with the headline column at 1.89 and 0.140. Employment remains near zero throughout.

Table A8: Anderson–Rubin weak-instrument-robust confidence intervals (headline exposure)

Outcome	2SLS $\hat{\beta}$	AR 95% CI
Skill mixing index ($\times$100)	1.89	$[1.44,\ 2.27]$
Any mixed skill (0/1)	0.032	$[0.029,\ 0.035]$
Mixed-skill count (intensive)	0.371	$[0.328,\ 0.414]$
Hourly wage (log)	0.140	$[0.128,\ 0.152]$
Employment rate (ACS)	0.026	$[-0.138,\ 0.189]$

Notes: This table reports Anderson–Rubin 95 percent intervals for the headline of Table 3, obtained by inverting the AR statistic after partialling the headline controls and fixed effects out of the outcome, regressor, and instrument (Frisch–Waugh–Lovell). Coefficients and interval endpoints are per one standard deviation of exposure (19.0 patent-units, 2010-employment-weighted). AR coverage is invariant to this rescaling. The inversion uses commuting-zone clustering. The Wald standard errors in Table 3 use two-way commuting-zone and state clustering. The annual flow benchmark appears in Appendix Table A11.

A.7 Shift-Share IV Diagnostics

The baseline design compares cells with different 2010 occupation shares, accounting for permanent cell differences and common demographic, state, and year shocks. The following diagnostics test whether the remaining variation is strong, well-timed, and distributed across shocks. Anderson Rubin intervals exclude zero for all posting outcomes and wages, but include zero for employment, so the main inference does not rely on a strong instrument. Table A11 varies the stock depreciation rate around the baseline $\delta = 0.10$ and reports the annual-flow alternative, including the ACS employment rate. The $\delta = 0.10$ column recovers the headline. Stock estimates remain positive and significant for mixing and wages across depreciation rates. The annual flow design gives a smaller and less precise mixing estimate because breakthrough classification needs forward citations, and most filings end after 2015.

Shock-level inference. The baseline already reports two-way cell-level standard errors and a shock-clustered effective F. One may still worry that many demographic observations inherit variation from the same 38 occupation shocks and therefore contain less independent information than their number suggests. Table A9 reports shock-level standard errors on the same column-(5) coefficients of Table 3. The point estimate, sample, weights, and fixed effects are unchanged. Only the unit used to measure uncertainty changes. The similarity of the conclusions under the two calculations makes duplication of cell observations an unlikely explanation for the main inference.

There may still be concerns about measurement before the main period. Lightcast coverage was limited in 2010 and improved in 2011. Table A10 maintains parser coverage at its 2011 level and reports a 2012–2013 placebo. The coverage-corrected placebo is -0.06 with a

Table A9: Shock-Level Standard Errors on the Headline Estimates

	$\widehat{\beta}$ (headline)	Two-way SE (CZ $\times$ state)	Shock-level SE	Shock-level 95% CI
Skill mixing index ($\times$100)	1.89***	[0.21]	[0.49]	[0.90, 2.88]
Prob. any mixed skill	0.032***	[0.002]	[0.008]	[0.016, 0.047]
Log # mixed skills (intensive)	0.371***	[0.032]	[0.096]	[0.177, 0.565]
Log hourly wage	0.140***	[0.010]	[0.034]	[0.071, 0.209]
Employment rate (ACS)	0.026	[0.093]	[0.125]	[-0.226, 0.279]

Notes: This table reprints the headline 2SLS estimates of Table 3, column (5), per 19.0 patent-units of exposure. Two-way standard errors are clustered by commuting zone and state. Shock-level standard errors treat the 38 occupation shocks as the independent units, clustering on the ISCO-2 occupation that carries each cell's largest 2010 employment share. They are computed from the same cell-level 2SLS, so only the standard error changes. The 95 percent interval uses the shock-level standard error and a t critical value with shock-cluster degrees of freedom. Sample, instrument (year-demeaned California leave-out instrument), weights, and fixed effects are those of column (5). The employment-rate row uses its own 411,784-cell ACS sample. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ (two-way).

standard error of 0.21. The coverage-corrected baseline is 2.13 with a standard error of 0.20, slightly higher than the raw estimate. This indicates that parser fill-in reduces the baseline effect rather than causing it.

Leave-one-ISCO-out. The baseline is not driven by any single 1-digit occupational group. Table A12 re-estimates the headline 2SLS specification of Table 3—all five baseline outcomes in parallel—after subtracting each ISCO-1 group's 2010 share contribution from the California leave-out instrument. The baseline row reproduces column (5) of Table 3. On the posting margins, the skill mixing index and the detection margins (any mixed skill, intensive log count) remain positive in every drop, and the index stays within 1.31–2.18 per SD—significant at the one percent level in eight of the nine drops and at five percent when ISCO 3 (Technicians) is excluded, the drop that weakens the instrument most and roughly triples the standard errors across the board. The log-wage coefficient is positive and significant at the one-percent level in all nine drops. The employment-rate coefficient is imprecise throughout, consistent with the precisely-estimated null in the baseline specification.

A.8 Stratified Recentering and Five-Category Mixing

The baseline centers the California shock across occupations within each year, removing the annual component common to all occupations. One concern is that comparing broad groups like professional, technical, service, and manual jobs may mix exposure with persistent wage differences. Table A13 addresses this by centering within each ISCO 1 group and year. It replaces $g^{\mathrm{CA}}_{ot} - \bar{g}^{\mathrm{CA}}_{t}$ with $g^{\mathrm{CA}}_{ot} - \bar{g}^{\mathrm{CA}}_{\iota(o),t}$, where $\iota(o)$ is the occupation's major group. All shares,

Table A10: Pre-Trend Placebos and Shock-Level Conditional Trend Test

	Pre-trend, coverage-corrected
Skill mixing index ($\times$100)	-0.062
	[0.212]
Log # mixed skills (intensive)	-0.005
	[0.010]
Log hourly wage	-0.010*
	[0.005]
Employment rate (ACS)	0.004
	[0.303]
Pre-window	2012–13
2012-lead test, wild-bootstrap p	0.889
Cells	31,518
Panel B. Shock-level conditional pre-trend test	
2011–2013 mixing-index change, conditional on 2010 level	-2.36
	[3.19]
p-value	0.464

Notes: Each entry is a 2SLS of the pre-period outcome change on post-2013 BKGPT exposure, instrumented by the California-recentered instrument and scaled by one SD (19.0 patent-units). Parser coverage is held at its 2011 level, so the window is 2012–2013. Specifications include commuting-zone and demographic-cell fixed effects and the 2010 covariates; standard errors are clustered on the ISCO-2 shock (38 clusters). Panel B regresses the 2011–2013 mixing-index change on the recentered exposure shock at the occupation level, conditional on the 2010 level. *, **, ***: 10, 5, 1 percent.

Table A11: Exposure stock and timing sensitivity

	$\delta = 0.05$	$\delta = 0.10$	$\delta = 0.15$	Annual flow
Skill mixing index $\times$100	1.08***	1.89***	2.80***	0.76*
	[0.14]	[0.21]	[0.30]	[0.42]
Log hourly wage	0.111***	0.140***	0.168***	0.061***
	[0.008]	[0.010]	[0.011]	[0.003]
Employment rate (ACS)	0.021	0.026	0.033	0.248**
	[0.075]	[0.093]	[0.110]	[0.119]

Notes: Columns (1)–(3) rebuild the cumulative breakthrough-GPT stock and its California year-demeaned leave-out instrument at the indicated depreciation rate from the continuous patent–occupation flows, then re-estimate the column-(5) specification of Table 3. The column δ=0.10 recovers that headline. Each stock column is per 1 SD of that column's 2010-employment-weighted exposure stock. The annual-flow column uses year-demeaned California flow exposure on the same specification, scaled by its own SD of 2.70 patent-units per year (matching Appendix Table **??**, column 4). The ACS employment-rate sample is 411,784. Standard errors two-way clustered by commuting zone and state in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

controls, fixed effects, and clustering stay the same. Randomization inference shuffles each occupation's full shock path within these groups 1,000 times, keeping the serial structure of the exposure path.

Table A12: Leave-One-ISCO-Out Robustness of the Shift-Share IV (Headline Exposure)

	Mixing index	Any mixed	Log count (int.)	Log wage	Emp. rate
Headline (full instrument)	1.89***	0.032***	0.371***	0.140***	0.026
	[0.21]	[0.002]	[0.032]	[0.010]	[0.093]
Drop ISCO 1: Managers	1.56***	0.041***	0.337***	0.133***	0.658*
	[0.36]	[0.007]	[0.053]	[0.010]	[0.365]
Drop ISCO 2: Professionals	1.52***	0.040***	0.412***	0.147***	-0.485
	[0.23]	[0.010]	[0.082]	[0.011]	[2.019]
Drop ISCO 3: Technicians	1.31**	0.106	0.665	0.333***	4.648
	[0.66]	[0.079]	[0.485]	[0.079]	[3.918]
Drop ISCO 4: Clerical support	1.91***	0.021**	0.333***	0.132***	-0.558
	[0.26]	[0.010]	[0.060]	[0.013]	[0.348]
Drop ISCO 5: Services and sales	2.07***	0.016	0.312***	0.124***	-0.040
	[0.29]	[0.020]	[0.117]	[0.017]	[0.967]
Drop ISCO 6: Skilled agricultural	1.90***	0.028***	0.348***	0.129***	0.387
	[0.19]	[0.002]	[0.035]	[0.008]	[0.461]
Drop ISCO 7: Craft and trades	1.96***	0.033***	0.371***	0.142***	0.307
	[0.21]	[0.003]	[0.040]	[0.011]	[0.600]
Drop ISCO 8: Operators and assemblers	2.04***	0.031***	0.399***	0.136***	-0.612
	[0.25]	[0.006]	[0.057]	[0.012]	[0.442]
Drop ISCO 9: Elementary occupations	2.18***	0.018	0.361***	0.107***	-1.244*
	[0.26]	[0.012]	[0.070]	[0.015]	[0.705]

Notes: This table re-estimates the headline 2SLS specification of Table 3. The endogenous regressor is B^{stock}_{ckt} (BKGPT exposure). The excluded instrument is rebuilt by subtracting the indicated ISCO-1 group's 2010 share contribution from the headline California leave-out instrument ($B^{\text{ca,stock,lo1}}_{ckt}$). Columns report the five headline outcomes in the same units as Table 3 (mixing index $\times$100). Coefficients are per one standard deviation of exposure (19.0 patent-units, 2010-employment-weighted). Standard errors in brackets are two-way clustered by commuting zone and state. The headline row reproduces column (5) of Table 3. Each of the nine rows drops one ISCO-1 group from the instrument. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The mixing index drops from 1.89 to 1.48 points per standard deviation and stays significant under stratified randomization inference with $p = 0.005$. The other two mixing margins have p-values of 0.006 or less. Employment remains close to zero. Wages decrease from 0.140 to 0.056, and the randomization inference value increases to 0.084. This means the main change in job content is still seen within broad occupation groups. The wage effect depends more on comparisons across groups, so I view the conditional estimate as a useful lower bound, not a replacement for the main design.

Column (5) of the same table keeps unclassified skills as a fifth content dimension when forming the index, so no skill is dropped and none is forced into a named category. The coefficient rises from 1.89 to 2.19 points per standard deviation (SE 0.19). Excluding the residual therefore attenuates, rather than generates, the estimated mixing response.

Table A13: Stratified Recentering and Five-Category Mixing

	Baseline (year-demeaned)		Stratified (ISCO-1×year)		Five-category
	2SLS (per SD)	RI p	2SLS (per SD)	RI p	2SLS (per SD)
	(1)	(2)	(3)	(4)	(5)
Panel A. Skill mixing in job postings					
Skill mixing index	1.89***	0.078	1.48***	0.005	2.19***
	[0.21]		[0.17]		[0.19]
Prob. any mixed skill (extensive)	0.0315***	0.026	0.0207***	0.006	—
	[0.0022]		[0.0017]		
Log # mixed skills (intensive)	0.371***	0.001	0.301***	0.004	—
	[0.032]		[0.025]		
Panel B. Local labor-market outcomes					
Log hourly wage	0.140***	0.008	0.056***	0.084	—
	[0.010]		[0.007]		
Employment rate (ACS)	0.026	0.908	-0.037	0.826	—
	[0.093]		[0.079]		
Effective F (shock-level cluster)	58.6		101.5		58.6
Recentering strata	year		year × ISCO-1		year
Randomization inference	within year		within ISCO-1		—
Observations	481,073 (postings); 411,784 (employment)				

Notes: This table reports just-identified BH-CA-2SLS coefficients on cumulative breakthrough-GPT exposure per one standard deviation (19.0 patent-units, 2010-employment-weighted), from a separate regression with the commuting-zone covariate vector and the commuting-zone, cell×year, and state×year fixed effects of Table 3. Columns (1)–(2) use the baseline year-demeaned California instrument (equation 4). Columns (3)–(4) recenter the same instrument within one-digit (ISCO-1) occupation groups, so identification uses only within-ISCO-1 variation in the shock. Column (5) re-estimates the column-(1) specification with the five-category skill mixing index that keeps unclassified skills as a residual content dimension; other outcomes are unchanged by that assignment and are marked —. Bracketed terms are standard errors two-way clustered by commuting zone and state. Stars are from those cluster-robust p-values. The RI p columns report reduced-form randomization-inference p-values from $R = 1{,}000$ permutations of the occupation shocks, permuted across all ISCO-2 occupations within a year in column (2), and within ISCO-1 groups (3–6 occupations each) in column (4). The effective F re-estimates the first stage with standard errors clustered at the shock level (38 ISCO-2 clusters). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ (cluster-robust).

B THEORY AND QUANTITATIVE

B.1 Propositions and Proofs

This appendix establishes the job design results used in Section IV. Throughout, $\mathbf{x} \in \mathbb{R}^K_{++}$, $\sigma \in (0,1)$, $\phi > 1$, and $\omega \in (0,1)$. The restrictions guarantee a well defined interior design. They also imply $\phi > \sigma$.

Proposition 1 (Optimal design and mixing)**.** *Let*

$$\Pi(\mathbf{y}) = (1-\omega) f(\mathbf{x}, \mathbf{y}) - \sum_{k=1}^{K} y_k^{\phi}$$

with f given by equation (6). The problem has a unique interior solution

$$y_k^* = \Lambda x_k^c, \quad c = \tfrac{\sigma}{\phi-\sigma}, \quad \Lambda = \left[\tfrac{1-\omega}{\phi}\,\Theta(\mathbf{x})^{1-\sigma}\right]^{1/(\phi-1)}, \quad \Theta(\mathbf{x}) = \left[\sum_k x_k^{(1+c)\sigma}\right]^{1/\sigma}. \tag{8}$$

At this solution

$$Mix(\mathbf{y}^*) = \frac{\sum_k x_k^c}{\sqrt{K}\sqrt{\sum_k x_k^{2c}}} \equiv M(\mathbf{x}, c). \tag{9}$$

The index belongs to $(0,1]$ and equals one exactly when the skill vector is balanced.

Proof. For $0 < \sigma < 1$, the CES aggregator is concave and homogeneous of degree one on the positive orthant. The design cost is strictly convex. Hence Π is strictly concave. Moreover,

$$f(\mathbf{x}, \mathbf{y}) \le K^{1/\sigma} (\max_k x_k) \|\mathbf{y}\|_\infty, \qquad \sum_k y_k^{\phi} \ge \|\mathbf{y}\|_\infty^{\phi}.$$

Since $\phi > 1$, profit tends to minus infinity as the design norm grows. A unique maximizer therefore exists.

The maximizer is not zero because $\Pi(t\mathbf{1}) > 0$ for sufficiently small $t > 0$. It is interior as well. If $y_k = 0$ at a nonzero maximizer, raising that coordinate by t changes output by a positive term of order t^{σ} and changes cost by t^{ϕ}. The output term dominates as t approaches zero because $\sigma < \phi$.

The first order condition is

$$(1-\omega) f^{1-\sigma} x_k^{\sigma} y_k^{\sigma-1} = \phi y_k^{\phi-1}.$$

Thus $y_k^{\phi-\sigma}$ is proportional to x_k^{σ}, which gives equation (8). Substitution into production gives $f(\mathbf{x}, \mathbf{y}^*) = \Lambda\Theta(\mathbf{x})$ and determines Λ. The cosine is homogeneous of degree zero, so its scale cancels and equation (9) follows. Cauchy Schwarz gives the upper bound and its equality condition. □

The exponent c is the elasticity of a relative requirement with respect to a relative worker skill. It is the single object through which σ and ϕ determine the shape of a design. Levels still depend on the two parameters separately.

Proposition 2 (Determinants of skill mixing)**.** *Suppose the worker skill vector is not balanced.*

(i) Mixing rises when σ falls.

(ii) Mixing rises when ϕ rises.

(iii) Let $K = 3$ and

$$c_* = \frac{1 - \sqrt{2/3}}{2}.$$

For $c \in [c_*, 1]$, every mean preserving contraction of one worker's three dimensional skill vector raises the mixing of the job designed for that worker.

Proof. Write $\kappa(t) = \log \sum_k x_k^t$. Equation (9) gives

$$\frac{d \log M}{dc} = \kappa'(c) - \kappa'(2c) = -\int_c^{2c} \kappa''(t)dt < 0, \qquad \kappa''(t) = \mathrm{Var}_{p_t}(\log x),$$

where $p_{t,k}$ is proportional to x_k^t. The inequality is strict away from a balanced skill vector. Since

$$\frac{\partial c}{\partial \sigma} = \frac{\phi}{(\phi - \sigma)^2} > 0, \qquad \frac{\partial c}{\partial \phi} = -\frac{\sigma}{(\phi - \sigma)^2} < 0,$$

parts (i) and (ii) follow.

For part (iii), it is enough to verify the Schur Ostrowski criterion for $h = \log M$. Let $S_1 = \sum_k x_k^c$ and $S_2 = \sum_k x_k^{2c}$. For two coordinates with $x_i > x_j$, set $A = x_i^c$, $B = x_j^c$, $r = B/A$, and $\rho = 1/c$. The criterion reduces to

$$\frac{A^{1-\rho} - B^{1-\rho}}{S_1} \leq \frac{A^{2-\rho} - B^{2-\rho}}{S_2}. \tag{10}$$

If $1/2 \leq c \leq 1$, the left side is nonpositive and the right side is nonnegative. Now consider $c < 1/2$. Minimizing S_2/S_1 over the third coordinate reduces equation (10) to

$$\psi_\mu(r) \leq Q(r), \qquad \psi_\mu(r) = \frac{r(1 - r^\mu)}{1 - r^{\mu+1}}, \qquad Q(r) = 2\left[\sqrt{(1+r)^2 + 1 + r^2} - (1+r)\right],$$

where $\mu = \rho - 2$. For $1/3 < c < 1/2$, convexity of $r^{\mu+1}$ gives $\psi_\mu(r) < \mu/(\mu+1) < 1/2$, while $Q(r) \geq 2/(1+\sqrt{3}) > 1/2$. For $c \leq 1/3$, let $\lambda = \mu/(\mu+1)$. Direct rearrangement gives

$$\psi_\mu(r) \leq \lambda \frac{2r}{1+r}.$$

The condition $c \geq c_*$ is equivalent to $\lambda \leq Q(1)$. The value $Q(r)$ is the positive root of $q^2/4 + (1+r)q - (1+r^2) = 0$. Evaluating this quadratic at $2\lambda r/(1+r)$ shows that the desired bound follows from

$$\lambda^2 r^2 + 2\lambda r(1+r)^2 - (1+r^2)(1+r)^2 \leq 0.$$

Indeed, $\lambda \leq Q(1)$ implies $\lambda^2 \leq 8(1-\lambda)$. After this substitution the negative of the left side is bounded below by

$$G(r) = r^4 + (2-2\lambda)r^3 + (4\lambda - 6)r^2 + (2-2\lambda)r + 1.$$

Writing $y = r + r^{-1}$ gives $G(r)/r^2 = (y-2)(y+4-2\lambda) > 0$ for $0 < r < 1$. Thus the Schur Ostrowski inequality holds throughout $[c_*, 1]$. Strictness follows along every nontrivial averaging of two unequal coordinates. □

The calibration uses $K = 3$. Its four values of c lie between 0.269 and 0.343, inside the sufficient interval. Part (iii) is a worker level statement. Aggregate occupational mixing also changes through employment weights, so its response to worker skill supply is computed rather than signed analytically.

B.2 Steady State Matching Equilibrium

Submarkets are indexed by worker type $\mathbf{x}$ and occupation o. A firm in a submarket chooses the design from Proposition 1. Meetings follow the constant returns matching function $M(s, v) = \mu s^\eta v^{1-\eta}$. Tightness is $\theta = v/s$, the job finding rate is $p(\theta) = \mu\theta^{1-\eta}$, and the vacancy filling rate is $q(\theta) = \mu\theta^{-\eta}$.

Let $r = (1-\beta)/\beta$. The worker and firm values satisfy

$$W(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}) = \frac{\omega f(\mathbf{x}, \mathbf{y}) - rU(\mathbf{x})}{r + \delta}, \tag{11}$$

$$J(\mathbf{x}, \mathbf{y}) = \frac{(1-\omega) f(\mathbf{x}, \mathbf{y}) - C(\mathbf{y})}{r + \delta}, \tag{12}$$

$$rU(\mathbf{x}) = b + \max_o p(\theta_{xo})[W(\mathbf{x}, o) - U(\mathbf{x})]. \tag{13}$$

Free entry gives

$$\kappa = q(\theta_{xo}) J_o(\mathbf{x}), \qquad \theta_{xo} = \left[\frac{\mu \Pi_o^*(\mathbf{x})}{\kappa(r+\delta)} \right]^{1/\eta}, \tag{14}$$

where Π_o^* is maximized flow profit. The numerical model evaluates this maximum on the fixed job bundle lattice.

The equilibrium is unique on the finite type grid under the maintained tie breaking rule. Firm values and tightness are explicit once the design is known. For each worker type, the right side of equation (13) is a continuous nonincreasing function of rU, while its left side is strictly increasing. A unique value follows. Given an irreducible skill transition matrix and positive separation, the induced finite Markov chain has a unique invariant distribution. Aggregate outcomes are computed from this distribution. These aggregation weights depend on the production parameters, which is why worker level design results do not by themselves identify aggregate responses.

B.3 External Calibration and Functional Forms

This subsection presents the calibration components derived from external sources, including search, matching, and discount parameters from the literature. It specifies the functional forms for production, costs, and matching, and describes the Markov process governing worker skill supply across two periods. Posting costs are calibrated to align with the two vacancy-to-unemployment ratios (see Table B1). P_1 and P_2 denote the two steady states, corresponding to the NLSY-79/97 waves ($P_1 = 2010$–12, $P_2 = 2021$–23). The Lightcast skill-mixing moment, which identifies the technology shift, incorporates postings up to the most recent year available. Section V.A provides details on the internally estimated parameters, and Appendix B.5 addresses their identification.

Sample. The model utilizes the same combination of NLSY79/97 and Lightcast data described in Section V to estimate three skill dimensions. NLSY provides information on worker abilities ($\mathbf{x}$) and tracks employment and wage changes, while Lightcast supplies occupational skill requirements ($\mathbf{y}$). The sample is restricted to the $P_1 = 2010$–12 and $P_2 = 2021$–23 NLSY waves and includes only workers with available skill data. Skill categories correspond to those in Sections I–III, except that analytical and computer skills are combined, resulting in a three-dimensional analysis ($K = 3$, $k \in \{$analytical/computer (a), interpersonal (p), routine (r)$\}$).

Grids and types. Occupations are categorized into high- or low-wage groups, as described in Section IV. High-wage jobs encompass managerial, professional, and white-collar roles, whereas low-wage jobs include blue-collar and service occupations. Each occupation's skill requirement y_k is mapped onto a grid, with one point representing 50 percent of the average observed y_k value.[48] Workers are classified as high- or low-type for each skill: individuals above the cross-sectional mean of x_k are pooled at the mean of above-mean values, and those below at the mean of below-mean values. With three skills, this classification yields eight worker types.

Functional forms. The multi-dimensional skill production function in equation (6) is informed by the multi-dimensional matching literature (Lise and Postel-Vinay 2020; Lindenlaub 2017; Ocampo 2022). It incorporates a CES complementarity parameter σ^j, which may differ across occupations. Job-design cost is specified as a convex function: $C(\mathbf{y}^j) = \sum_{k=1}^{K} (y_k^j)^{\phi^j}$, where ϕ^j determines the rate of cost increase for occupation j. The matching function adopts a Cobb–Douglas form: $M(s, v) = \mu s^\eta v^{1-\eta}$, with η representing the elasticity of matches with respect to search effort and μ denoting matching efficiency. This formulation yields a job-finding rate $p(\theta) = \mu\theta^{1-\eta}$ and a vacancy-filling rate $q(\theta) = \mu\theta^{-\eta}$.

External parameter values. The model operates on a one-year period. For risk-neutral agents, $\beta = 0.96$ reflects a 4 percent annual interest rate. The separation rate δ is set at 10 percent, as in Shimer (2005). The worker share of output ω is 0.6, consistent with the 2005 labor share of GDP. Unemployment benefits b are set at 41.5 percent of the earnings loss for the lowest-paid occupation, following Braxton, Herkenhoff, and Phillips (2020). Matching elasticity η is 0.5 and efficiency μ is 0.65, based on Mercan and Schoefer (2020). The vacancy-cost scale κ is 0.884 in P_1 and 0.724 in P_2, calibrated to match the two vacancy-to-unemployment ratios. Panel A of Table B1 presents these values.

Analytical/computer skill increases approximately 3.6 times faster than it depreciates (annual rates 0.36 versus 0.10; Table B1, Panel B).[49] Interpersonal skill evolves slowly in both directions, while routine skill adjusts most rapidly.

B.4 Calibration of Worker Skill Supply

The evolution between $P_1 = 2010–12$ and $P_2 = 2021–23$ is determined by workers' choices regarding jobs, education, and skill-building at work. This evolution is not estimated within the criterion but is calibrated externally from NLSY data, following Lise and Postel-Vinay (2020). Together with the calibrated posting-cost path, this forms the labor-supply component

[48] Grid points are anchored at the median of each occupation's two-period average, ensuring consistency across periods.

[49] Lise and Postel-Vinay (2020) report monthly estimates, which are annualized here.

Table B1: Externally calibrated parameters

Parameter	Description	Value	
Panel A. Externally calibrated: discounting, separation, and matching			
β	Annual discount factor	0.96	
δ	Exogenous separation rate	0.10	
ω	Worker share of match output (labor share)	0.60	
b	Unemployment benefit (share of lowest-occ. wage)	0.415	
η	Elasticity of the matching function	0.50	
μ	Matching efficiency	0.65	
κ	Vacancy-cost scale (P_1, P_2)	0.884	0.724
Panel B. Externally calibrated: annual skill adjustment speed		Up	Down
γ_a	Adjustment speed, analytical/computer skill	0.36	0.10
γ_p	Adjustment speed, interpersonal (social) skill	0.05	0.00
γ_r	Adjustment speed, routine skill	1.00	0.36

Notes: This table reports the parameters held outside the simulated-method-of-moments criterion. Panel A collects discounting, separation, bargaining, benefits, and matching frictions; Panel B collects skill-specific annual adjustment speeds. Separation follows Shimer (2005); the worker share matches the labor share of GDP in 2005; unemployment benefits follow Braxton, Herkenhoff, and Phillips (2020); matching elasticity and efficiency follow Mercan and Schoefer (2020). The vacancy-cost scale κ is set period by period to the values that reproduce each environment's vacancy-to-unemployment ratio (0.884 in P_1, 0.724 in P_2). Skill adjustment rates in Panel B are annualized from Lise and Postel-Vinay (2020), aligning analytical/computer, interpersonal, and routine skills with that paper's cognitive, interpersonal, and manual categories. The eight production parameters (σ, ϕ) are reported in Table 6. The high-skill bin shares that shift worker supply $G(\mathbf{x})$ between environments are stated in Appendix B.4.

in Figure 11.

Adjustment process. Worker skills evolve according to the gap between current skills and the requirements of their job or college major. Skills increase when requirements exceed initial skills and decrease when requirements are lower or during unemployment. The adjustment rate is skill-specific and varies depending on whether the skill is increasing or decreasing. Each year, a worker's skill k changes by γ_k times the gap between their skill and the job's requirement, with γ_k differing for increases and decreases. Unemployed workers may experience skill loss, as unemployment is modeled with zero requirements; however, skill levels cannot fall below their initial value. For college, workers are assumed to spend approximately three years acquiring the skills estimated by Lise and Postel-Vinay (2020). Two modifications are applied to Lise and Postel-Vinay (2020)'s estimates: weekly rates are annualized by multiplying by 47 working weeks, and the cognitive, interpersonal, and manual categories are mapped to analytical, interpersonal, and routine skills, respectively. Since Lise and Postel-Vinay (2020) do not estimate computer skill, the cognitive-skill estimate is used as a proxy. In calculating skill adjustment, both worker skills and job requirements are standardized. For instance, if a job requires one standard deviation more analytical skill than the worker possesses, the worker acquires 0.36 standard deviations of analytical skill in a year. If a worker's interpersonal skill is one standard deviation above the requirement, it decreases by only 3×10^{-5} standard deviations per year, which is nearly unchanged (see Table B1, Panel B).

Period-1 to period-2 shift, by skill. During the calibrated period, $G(\mathbf{x})$ shifts moderately toward a greater proportion of high-skill workers. The share of workers in the high-analytical/computer group increases from 0.49 to 0.60 (+0.114), the high-interpersonal share rises from 0.48 to 0.55 (+0.071), and the high-routine share remains nearly unchanged (0.47 to 0.46, −0.004). The majority of the change in period-2 skill supply is concentrated in the analytical/computer category, consistent with cohort and education trends identified by Lise and Postel-Vinay (2020) and the NLSY data for these periods. These externally calibrated shifts, together with posting costs, constitute the labor-supply component in Figure 11. The production parameters account for the remaining changes after this supply shift is incorporated into equilibrium.

B.5 Internal Estimation, Identification, and Inference

This subsection details the simulated-method-of-moments criterion employed to estimate the production-side parameters and outlines the identification strategy. The estimation utilizes responses from Section III, measured as levels in each calibrated setting. Identification is challenging because, with multiple skills, a task vector in the model, and a CES production function, simple moments such as wage levels cannot distinguish the CES substitution parameter σ from the cost function's curvature ϕ (Caselli and Coleman, 2006). This approach leverages the shape of the skill bundle, as well as relative prices and employment, to address this identification issue.

Estimator. Let $\theta = \left(\sigma_o^{P_t}, \phi_o^{P_t}\right)_{o\in\{1,2\},\, t\in\{1,2\}}$ denote the eight production parameters — a CES substitution parameter and a cost-convexity parameter in each occupation in each period. Let $\widehat{m}$ stack the ten targeted levels — the two projected mixing responses, the occupational wage-gap response, the high-wage employment share, and unemployment, each evaluated in both environments — and $m(\theta)$ their model counterparts. The estimator solves

$$\widehat{\theta} = \arg\min_{\theta\in\Theta} \sum_j w_j \left(\frac{m_j(\theta) - \widehat{m}_j}{s_j}\right)^2 + \Pi(\theta). \tag{15}$$

The scales s_j standardize wages, shares, and mixing for comparison. The weights w_j assign importance to each moment and are determined prior to the main search. $\Pi(\theta)$ is a soft penalty that discourages increases in either substitution parameter and does not bind at the promoted point. The setup requires an interior, nondegenerate equilibrium, greater aggregate mixing, and a consistent order of the production response. Each candidate parameter vector solves both steady states on the same grid with identical simulation draws, ensuring that any differences in the criterion arise from parameter changes rather than simulation noise.

Targeted moments. The estimation targets ten levels (see Appendix Table B2). Three measures from Section III—the low-wage mixing response, the high-wage mixing response, and the occupational wage-gap response—are included for each period's supply and search conditions. High-wage employment and unemployment serve as anchors for each environment. Consequently, the ten rows in the criterion represent fewer than ten independent facts: Section III estimates are pooled across periods, and repeated response rows assess whether a single technology change produces similar effects in two labor markets. As these repetitions do not introduce new sampling variation, a standard overidentification test is not reported.

Identification of $\sigma_o^{P_t}$. The mixing responses identify the substitution parameters. When complementarity increases, firms adjust skill requirements more in response to shocks. Occupations with a narrower initial skill bundle require greater rebalancing. Since low-wage occupations are more specialized than high-wage ones, comparing their response magnitudes helps distinguish the two σ changes.

The wage gap, together with the anchors, assists in identifying the cost side. Given the mixing response, the share of surplus from redesign allocated to the low-wage occupation depends on the cost of the new bundle, so the gap response informs the determination of (ϕ_1, ϕ_2). This approach is effective where wage levels are not, since levels change with efficiency normalization, but the gap remains stable. The high-wage employment share and unemployment serve as anchors for the environment affected by the shock, but are not the primary identifiers for σ.

Fit. Table B2 presents each targeted level alongside its model value. Eight of ten conditions are within twenty-five percent of their targets. The model matches the wage-gap levels exactly and closely approximates low-wage mixing; the largest discrepancies occur in the high-wage mixing responses. This pattern supports the identification argument that the criterion captures the joint change in bundle shape and the ordering led by low-wage occupations more distinctly than the smaller high-wage response.

Table B2: Moment inventory: the ten levels and their fit

Moment	Target	Model
Period 1 (2010–12)		
Mixing index, low-wage occ.	0.930	0.958
Mixing index, high-wage occ.	0.968	0.969
Occupational wage gap (high/low)	1.385	1.385
High-wage employment share	0.607	0.675
Unemployment rate	0.109	0.089
Period 2 (2021–23)		
Mixing index, low-wage occ.	0.940	0.960
Mixing index, high-wage occ.	0.975	0.977
Occupational wage gap (high/low)	1.512	1.512
High-wage employment share	0.646	0.694
Unemployment rate	0.058	0.065

Notes: This table reports each moment as a level measured at the state the economy is observed in, not a response to a shock. The mixing indices are occupation-by-era means of the named-skill mixing measure of Section I, mapped into the model's band by the frozen slope 0.25576. The wage gap is the ratio of mean earnings between the two occupation groups. The share and the unemployment rate are cross-sectional labour-market anchors.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Altonji, J. G., Bharadwaj, P., and Lange, F. (2012). Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4):783–828.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China syndrome: local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168.

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

Autor, D. H. and Price, B. (2013). The changing task composition of the us labor market: An update of autor, levy, and murnane (2003). *Working Paper*.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213.

Braxton, J. C., Herkenhoff, K. F., and Phillips, G. M. (2020). Can the unemployed borrow? implications for public insurance. Technical report, National Bureau of Economic Research.

Braxton, J. C. and Taska, B. (2023). Technological change and the consequences of job loss. *American Economic Review*, 113(2):279–316.

Caselli, F. and Coleman, Wilbur John, I. (2006). The world technology frontier. *American Economic Review*, 96(3):499–522.

De Chaisemartin, C. and d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, pages 1–45.

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

Goldin, C. and Katz, L. F. (2010). *The race between education and technology*. harvard university press.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools.

Hansen, L. P., Heaton, J., and Yaron, A. (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics*, 14(3):262–280.

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

Kim, G., Merritt, C., and Peri, G. (2024). Measuring and predicting "new work" in the united states: The role of local factors and global shocks. Technical report, National Bureau of Economic Research.

Lightcast (2024). Representativeness analysis of lightcast job posting data - u.s. Technical report, Lightcast.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, 84(2):718–789.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–76.

Mercan, Y. and Schoefer, B. (2020). Jobs and matches: Quits, replacement hiring, and vacancy chains and vacancy chains. *American Economic Review: Insights*, 2(1):101–124.

Ocampo, S. (2022). A task-based theory of occupations with multidimensional heterogeneity. *Working Paper*.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49.