Occupational Skill Mixing Under Breakthrough GPTs*

Elmer Zongyang Li†

January, 2026

Click here for the latest version

Abstract

Leveraging a near-universe of U.S. online job postings alongside task descriptors from 2005 to 2024, this paper documents a substantial increase in skill mixing—the integration of skills previously specific to other occupations. 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 non-routine computer and analytical skills. This rise is further corroborated when measured by an intensive margin of skill balancing. To explain these patterns, I examine the impact of breakthrough general-purpose technologies (GPTs)—innovations that are simultaneously highly novel and broadly applicable—identified using patent data and mapped to occupational usage. Event-study estimates show that exposure to a breakthrough GPT raises the number of mixed skills by about 4 percent on impact—an effect that persists—and increases wages and employment by 2.7 and 8.5 percent within six years, with effects concentrated in low-skill occupations. A calibrated multidimensional matching model with endogenous occupation design attributes the breakthrough GPT driven rise in mixing primarily to greater complementarity among skills and more convex occupational costs, which also contribute to aggregate occupation wage gaps and nearly half of the changing employment ratios.

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

*I am extremely grateful to Philipp Kircher, Michael Lovenheim, Julieta Caunedo, and Mathieu Taschereau-Dumouchel for their guidance and continuous support. I also want to thank Justin Bloesch, Xiaoming Cai, Maria Fitzpatrick, 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 combine a broader set of skill requirements, this would indicate stronger complementarity and could raise productivity and labor market returns.

This paper studies "skill mixing"—the tendency for employers to bundle together skill demands that were previously seen in different occupations—and examines its drivers and labor market consequences. I begin by constructing comprehensive data sources of employer skill demands. First, I leverage the near-universe of online job postings from Lightcast (formerly "Burning Glass") between 2010 and 2024 to capture real-time shifts in the extensive margin of skill demand. Lightcast systematically scrapes over 220,000 online sources, including job boards and corporate career pages, and parses vacancy text into a taxonomy of more than 33,000 skills, which allows me to track detailed skill demands.[1] Second, I complement Lightcast with the Occupational Information Network (O*NET), which provides stable, task-based measures of occupational requirements derived from job incumbent surveys. To quantify the intensive margin of these changes, I introduce a "mixing index" calculated as the cosine similarity between an occupation's skill vector and a perfectly balanced unit vector, where a higher index indicates a more balanced skill portfolio.

To characterize skill mixing in the U.S. labor market, I begin by decomposing changes in skill demand using the Lightcast data. I find that the composition of occupational skills has churned significantly: nearly half (47.6 percent) of the skills observed in 2024 were not present in the same occupations in 2010, and almost all of these "new" requirements are "mixed skills"–skill requirements that appeared in other occupations in 2010.[2] This pattern is especially pronounced in low-skill occupations, such as Operators and Elementary

[1]While Lightcast job postings are not a random sample of all vacancies, several recent studies document that they closely track administrative benchmarks of aggregate vacancy flows and cross-sectional occupational distributions (Lightcast 2024; Tsvetkova et al. 2024). A growing literature relies on the parsed skill information to document changing skill demands, including Deming and Kahn (2018), Hershbein and Kahn (2018), and Braxton and Taska (2023). Following Kim, Merritt, and Peri (2024), I reduce redundancy by merging roughly 2,000 skills that appear similar using cosine similarity of skill embeddings.

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

occupations.[3] The importance of skill mixing remains when measured by the share of postings that contain mixed skills and is robust to controls for occupational composition, regional composition, and labor supply.

Three stylized facts 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 computer and analytical skills, including categories such as IT and marketing, which are general-purpose in nature; by contrast, domain-specific and routine skills, such as architecture or facilities management, are underrepresented.[4] Third, the rise in skill mixing is corroborated by the intensive margin of skill balancing: the mixing index has also increased over time during this period, and this trend appears when using O*NET data as well.[5]

What drives the rise in skill mixing? I examine the impact of breakthrough general-purpose technologies (GPTs)—innovations that are simultaneously highly novel and broadly applicable. I construct a new measure of exposure to breakthrough GPTs by combining patent data from the USPTO PatentsView database with O*NET technology descriptors. Following the methodology of Kelly et al. (2021), I define a patent as a "breakthrough" if it is textually distinct from prior patents but highly influential for citing patents. I combine this with the generality index of Hall, Jaffe, and Trajtenberg (2001) to isolate technologies that impact a broad range of sectors. Using high-dimensional vector embeddings, I link these patents to specific O*NET technology descriptors to measure occupational exposure. The occupational composition of these patents reveals a distinct pattern: while standard breakthrough patents remain concentrated in high-skill roles like Professionals and Technicians, breakthrough GPTs show a broader occupational composition and are more low-skill oriented.

To estimate the causal impact of breakthrough GPT, I employ a dynamic difference-in-differences design with staggered adoption following De Chaisemartin and d'Haultfoeuille (2024).[6] I find that exposure to a breakthrough GPT leads to a sharp, persistent increase

[3] I use the ISCO occupational classification as it is the taxonomy provided in the Lightcast data. I show that the main patterns of skill mixing are robust when using the Census SOC classification.

[4] The classification follows two distinct approaches: (i) the standardized Lightcast Open Skills taxonomy, which groups skills into broad career areas such as "Information Technology" and "Finance"; (ii) Routine/Non-Routine (RNR) categories drawn from the task-based literature.

[5] To address O*NET's staggered updating schedule, I clean the data by focusing on broad intervals (2005–2024) during which occupations were updated at least twice. I further restrict the analysis to task descriptors derived exclusively from job incumbent surveys to ensure longitudinal consistency and avoid potential biases from analyst ratings.

[6] Recent econometric literature demonstrates that in staggered adoption designs with heterogeneous treatment effects, standard Two-Way Fixed Effects (TWFE) estimators can yield biased results due to the negative weighting of earlier-treated units. The De Chaisemartin and d'Haultfoeuille (2024) estimator addresses this by isolating

in skill mixing.[7] Specifically, the number of mixed skills required in an occupation rises by approximately 4 percent upon impact and an accumulated 11 percent in six years; the mixing index rises by an accumulated 1.6 points over the same horizon. This average effect, however, masks substantial heterogeneity. Splitting the sample by skill level reveals that for low-skill occupations, exposure leads to a sizable increase in the accumulated number of mixed skills (14 percent) and a significant rise in the mixing index (1.8 points) over a six-year period. In contrast, high-skill occupations experience a much smaller rise in the number of mixed skills (5 percent) and a statistically insignificant change in the mixing index. This result establishes breakthrough GPTs as the primary driver of the substantial skill mixing observed in lower-skill occupations.

These shifts translate into tangible economic returns that mirror the heterogeneity in skill mixing. Among low-skill occupations, exposure to breakthrough GPTs leads to a statistically significant 2.7 percent increase in hourly wages and an 8.5 percent increase in employment over six years. In contrast, the effects are smaller and statistically weaker for high-skill occupations. These findings reinforce the hierarchical skill demand diffusion: breakthrough GPTs allow high-value, non-routine tasks (such as IT and design) to be bundled into traditionally manual or routine jobs, effectively upskilling lower-skill roles. Unlike standard automating technological change that predicts displacement, these results suggest that breakthrough GPTs complement low-skill work, increasing the marginal product of labor and raising both pay and employment as employers expand the set of skill demands.

To provide a structural interpretation of the effects of breakthrough GPTs, I develop a multi-dimensional matching model in which firms choose their optimal skill requirements by balancing the production benefits of bundling skills against the convex costs of organizing complex tasks, taking into account the complementarity across different skills, the efficiency of each skill, and worker skill supply.[8] I calibrate the model to match the causal response of skill mixing to breakthrough GPTs and key moments of the U.S. labor market. Counterfactual analysis reveals that breakthrough GPTs' impact on skill mixing is overwhelmingly driven by

the treatment effect through comparisons of switchers to stable control units, while also accommodating the non-absorbing nature of GPT exposure, where occupations may be treated multiple times.

[7] I measure outcomes at the commuting zone-by-occupation level because it allows me to leverage within-zone variation in occupational composition while adjusting for broader local economic shocks. This level of aggregation captures equilibrium adjustments in local labor markets—where labor supply and demand interact—while permitting the inclusion of commuting zone-by-year and occupation-by-year fixed effects to isolate the specific impact of technological exposure.

[8] The model is embedded in a directed search environment where firms post contracts specifying both wages and skill requirements. This structure is critical for tractability: it renders the equilibrium block recursive, allowing me to solve for the firm's optimal skill demands and prices independent of the complex distribution of workers in the high-dimensional skill space. This property bypasses the "curse of dimensionality" typically found in matching models with heterogeneous agents (Menzio and Shi, 2010).

technological shifts. Higher skill complementarity accounts for 73 percent of the GPT-driven increase in skill mixing, and changes in occupational skill costs contribute 25 percent, while changes in skill efficiencies and shifts in worker skill supply play a minimal role. These two forces also drive the aggregate increase in the wage gap between high- and low-skill occupations, and nearly half of the employment gap during this period.

The rest of the paper is organized as follows. The next section reviews the literature and outlines contributions. Section II presents the data and descriptive facts on skill mixing. Section III establishes the causal role of breakthrough GPTs and their labor market effects. Section IV details the multi-dimensional matching and counterfactual analyses. Section V concludes.

I Literature Review

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). 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 toward a broader range of skills, where employers increasingly mix skills that were previously specific to other occupations. 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).[9]

My analysis also connects to the emerging research on skill demand shifts using real-time job posting data. Recent studies establish that skill requirements in job postings are heterogeneous and predictive of firm performance (Atalay and Sarada, 2020; Deming and Kahn, 2018), shift toward higher requirements during recessions (Hershbein and Kahn, 2018), and can generate earnings losses when displaced workers lack specific new skills (Braxton and Taska, 2023). This paper instead studies the mixing of skills driven by breakthrough GPTs. Complementing the view of upskilling as an increase in requirements or a source of displacement, I show that skill mixing represents a hierarchical broadening of job roles, where high-skill, non-routine skills are integrated within lower-skill occupations. I document that this bundling of complex skills, driven by breakthrough GPTs, generates innovational complementarities that raise wages and employment for low-skill workers.

[9]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.

This paper directly contributes to the literature analyzing the impact of GPTs. While early literature was largely theoretical (e.g., Helpman and Trajtenberg 1994), empirical evidence highlights that the diffusion of GPTs, such as the steam engine, is often a slow process constrained by the need for downstream sectors to adapt (e.g., Crafts 2004). Recent research emphasizes that GPTs raise productivity not only through direct effects but also by inducing costly organizational changes and complementary intangible investments (Brynjolfsson, Rock, and Syverson 2021; Juhász, Squicciarini, and Voigtländer 2024). In the context of Information and Communications Technology (ICT), this mechanism is considered to explain both the U.S. productivity advantage (Bloom, Sadun, and Reenen 2012) and the evolution of labor demand (Katz and Murphy, 1992; Autor, Levy, and Murnane, 2003). However, fewer studies cleanly measure exactly how job requirements themselves reorganize—specifically through the bundling or mixing of different skills—as GPTs diffuse. I provide direct evidence that breakthrough GPTs induce skills to be mixed from high-skill occupations into lower-skill jobs, and evaluate the resulting 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 trade-offs 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.

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

How have the skill demands of U.S. occupations evolved over the past decade? I begin by documenting a key feature of the U.S. labor market from 2010 to 2024: the increasing prevalence and intensity of skill mixing across occupations, using detailed vacancy-level data from Lightcast and supplemented with task-based occupation-level data from ONET. I first present evidence that the breadth of skills required within occupations in job postings has expanded over time, driven by cross-occupation skill mixing. To better understand the structure of skill mixing, I apply network analysis to visualize the flow and composition of skills that are mixed across occupations. Finally, I cross-validate these findings using ONET data, employing an angle-based vector similarity measure constructed from a smaller, harmonized set of skills. This confirms that the trend toward skill mixing is also reflected in established task-based measures.

II.A Skill Mixing in Job Postings

Data and Measurement: To analyze the evolution of 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 de-duplication 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.[10]

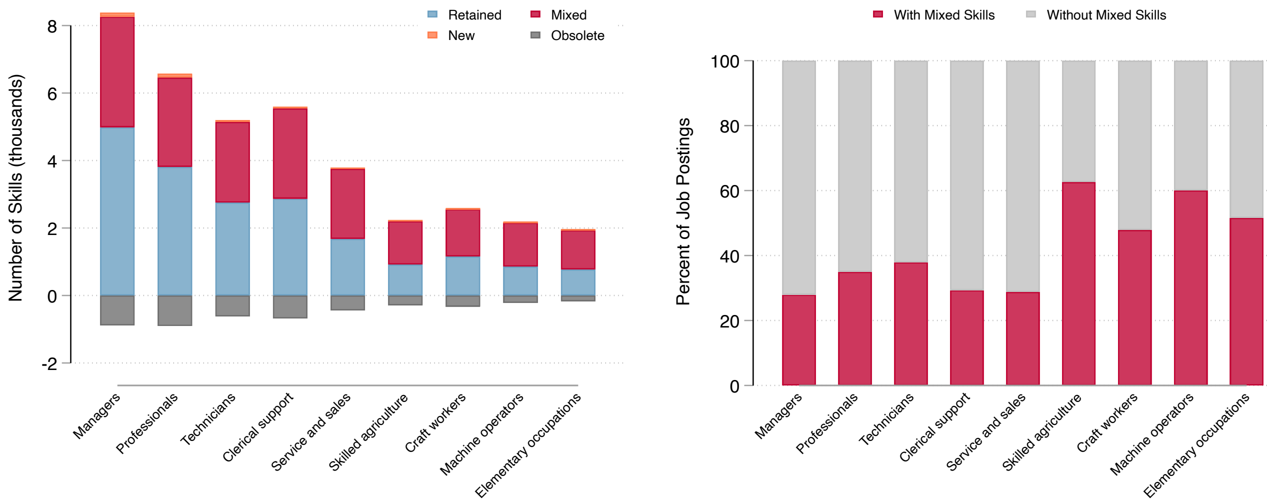
Importantly, Lightcast parses detailed skill requirements from raw vacancy text, and prior work uses this information to study trends in job skill demand (Deming and Kahn 2018; Hershbein and Kahn 2018; Braxton and Taska 2023). The database includes a taxonomy of over 33,000 continuously updated skills. 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).

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. Figure 1 illustrates this composition, classified at the 4-digit ISCO level and averaged within 1-digit occupational groups. 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 be-

[10] 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



(a) Composition of Skills

(b) Composition of Job Postings

Notes: Panel (a) presents the distribution of skill types across major occupational groups in the United States, distinguishing four categories of skills based on their emergence and persistence over time. The occupation classification follows International Standard Classification of Occupations (ISCO), with calculations first conducted at the 4-digit level and then averaged at the 1-digit level. Values are expressed in thousands, and declining skills are plotted below the horizontal axis for visual clarity. Panel (b) presents the share of job postings in each 1-digit ISCO group that require at least one mixed skill.

Table 1: Annual Changes in Skill Mixing

	Log number of mixed skills			Mixing posting share (%)		
	(1)	(2)	(3)	(4)	(5)	(6)
Year	0.15***	0.16***	0.16***	1.61***	1.63***	1.72***
	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]	[0.02]
Czone × Occ		✓	✓		✓	✓
Worker controls			✓			✓

Notes: This table reports estimates of the annual increase of skill mixing (captured by the variable "Year") using different occupational skill mixing measures. Columns (1)–(3) use the log number of mixed skills as the dependent variable; columns (4)–(6) use the percentage of postings that mix skills. Standard errors are clustered at the czone level. Statistical significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

tween 2010 and 2024 occurred through the mixing of existing skills across occupations. On average, 47.6 percent of the skills observed in occupations in 2024 were not present in the same occupation in 2010, and 45.4 percent of these newly observed skills were mixed in from other occupations. This pattern holds across all major occupational groups. While high-skill occupations demand the largest number of mixed skills, the share of mixing is higher in low-skill occupations.

Robustness checks: One concern is that a simple count of skills may not reflect their importance. Panel (b) addresses this by examining the percentage of postings that require at least one mixed skill. This measure reveals that skill mixing is widespread and disproportionately high in lower-skill occupations. Specifically, 48 to 62 percent of postings in Skilled Agriculture, Craft Workers, Machine Operators, and Elementary occupations require mixed skills, compared to 28 to 38 percent in high-skill occupations. Another concern is that 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 significant, averaging 49 percent for low-skill and 37 percent for high-skill occupations.

Further, one may worry that the upward trend reflects shifts in occupational composition, regional composition, or labor supply (for example, rising education or demographic change), rather than changing labor demand. Table 1 addresses this by estimating the annual time trend in skill mixing while adjusting for occupation groups and commuting zones, as well as for local labor market composition, which includes the local shares of female, racial minority, and college-educated workers, along with age structures and unemployment rates. Columns (1)–(3) use the log number of mixed skills; Columns (4)–(6) use the share of postings requiring at least one mixed skill. The baseline specifications imply a strong upward trend: mixed skills rise by about 0.15 log points per year and the mixed-posting share by about 1.61 percentage points per year. These estimates remain stable, and even slightly larger, when adding commuting zone-by-occupation fixed effects and the full set of worker controls, with the trend settling around 0.16 for the count measure and 1.72 percentage points for the posting share. This stability indicates that the rise in skill mixing is not merely driven by the composition of occupations, regions, or labor supply.

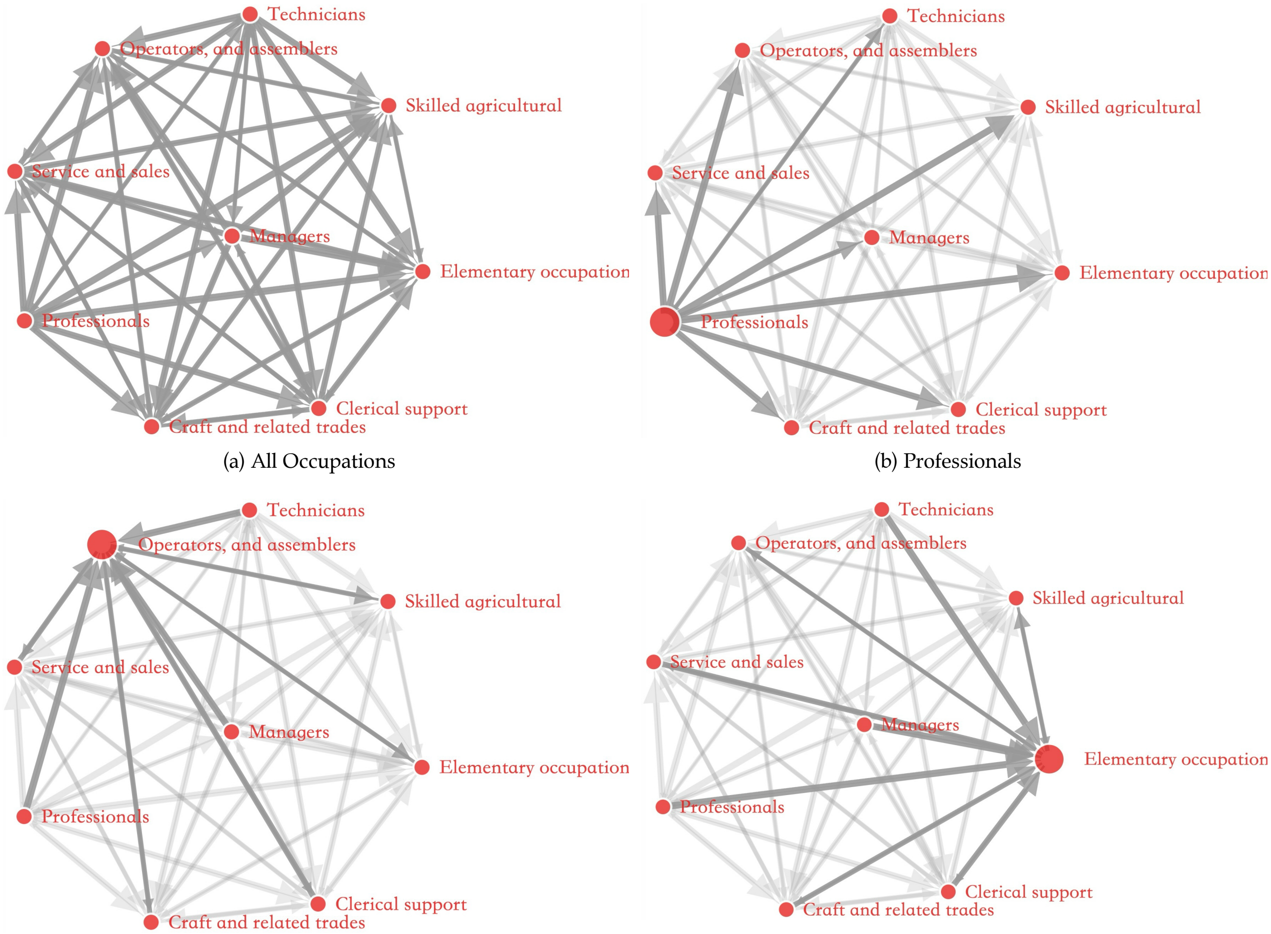
II.B Network and Content of Skill Mixing

To understand the skill mixing in the United States between 2010 and 2024, I conduct two analyses. First, I use a network analysis to examine the direction of skill flows across occupations, identifying which occupations serve as the main sources of skill mixing and which act as destinations. Second, I examine the composition of mixed skills within occupations across skill categories to examine the content of skill mixing.

Network of skill mixing: Figure 2 panel (a) first visualizes the flow of skill mixing across occupational groups between 2010 and 2024 using a force-directed network.[11] Each node

[11] 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

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



(a) All Occupations

(b) Professionals

(c) Operators and Assemblers

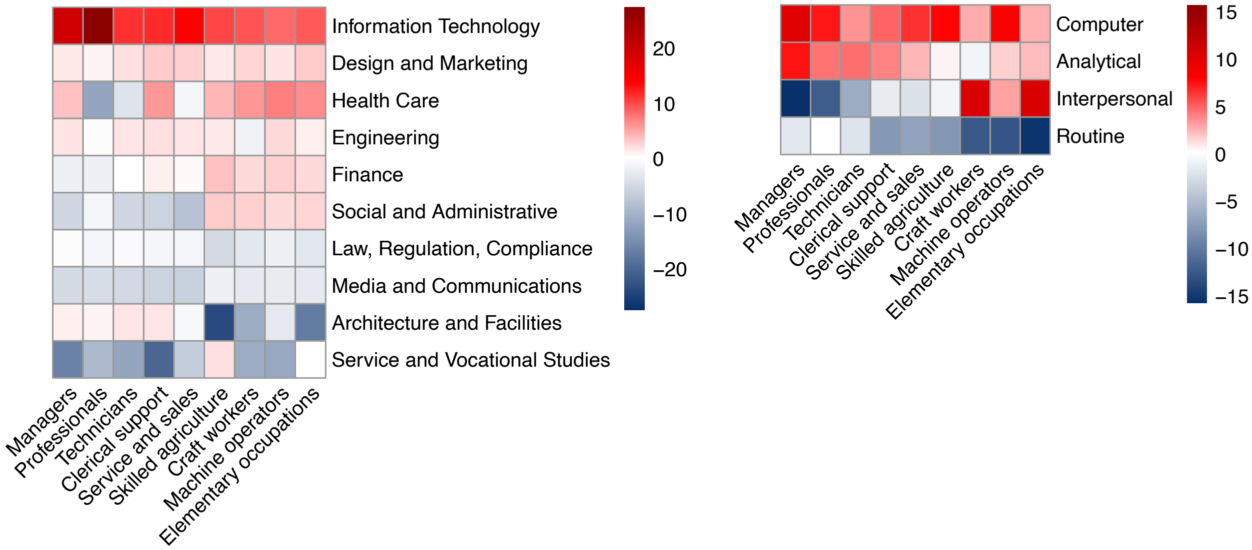
(d) Elementary Occupations

Notes: This figure visualizes the network structure of occupational skill mixing and the co-occurrence of five major skill categories within U.S. job postings from 2010 to 2024. Panel (a) displays a force-directed network of skill transfer across 4-digit ISCO occupations, where each node represents an occupation and edge weights reflect the number of shared new skills between occupations. Panels (b) and (c) depict co-occurrence networks of major skill categories in 2010 and 2024, where each node corresponds to a skill group and edges represent the share of job postings in which two skills appear together. The table shows how changes in the prevalence of mixed skills within 4-digit ISCO occupations relate to shifts in the intensity of major skill categories. Skill intensity is defined as the share of job postings within each 4-digit ISCO occupation that include the relevant keywords. All variables in the table are expressed as year-on-year differences, and the regression includes year and occupation fixed effects.

represents an ISCO one-digit occupation, and each directed edge captures the extent to which skills appearing in one occupation in 2010 also appear in others in 2024. The thickness of the edge corresponds to the count of skills. The network reveals that high-skill occupations such as Professionals and Technicians are key sources of skill mixing. In contrast, low-skill

equilibrium, so that occupations sharing more mixed skills are positioned closer, while those with fewer shared skills are pushed apart—naturally revealing clustering patterns.

Figure 3: Relative Skill Mixing Intensity Across Categories



Notes: This figure presents the distribution of skill types across major occupational groups in the United States, distinguishing four categories of skills based on their emergence and persistence over time. The occupation classification follows International Standard Classification of Occupations (ISCO), with calculations first conducted at the 4-digit level and then averaged at the 1-digit level. Values are expressed in thousands, and declining skills are plotted below the horizontal axis for visual clarity.

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: We next analyze the content of skill mixing by examining the specific categories of skills being mixed. Figure 3 presents a heatmap of the relative mixing intensity of different skill categories, defined as the difference between a category's share within the set of mixed skills and its share in the economy-wide stock of skills.[12] A positive

[12]Normalizing by the economy-wide skill stock filters out pure size effects across skill categories, ensuring

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.[13] Information Technology and Design and Marketing skills emerge as the two most highly mixed categories, exhibiting large positive values across different occupations. We also observe that low-skill occupations 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 onto four Routine/Non-Routine (RNR) groups—Analytical, Interpersonal, Computer, and Routine-following Acemoglu and Autor (2011).[14] I first compute vector embeddings for every Lightcast skill and assigning them to the category scoring the highest cosine similarity with the representative keywords from Braxton and Taska (2023) and Hershbein and Kahn (2018).[15] This classification confirms the technological bias shown in panel (a): Computer skills are the overwhelming driver of mixing across all occupations, followed by Analytical skills. Low-skill occupations also exhibit higher mixing of Interpersonal skills. In stark contrast, Routine skills exhibit strong negative values across all occupations.

Taking stock, the evidence shows that skills from knowledge-intensive occupations are being integrated into lower-skilled roles, thereby increasing their complexity. This shift is mainly driven by Computer and Analytical skills, such as those classified as IT and marketing-related, which are "general-purpose" in nature. These patterns highlight a clear directional trend: high-skill occupations are shaping the broader skill demand through the diffusion of broadly applicable, non-routine skills.[16]

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.

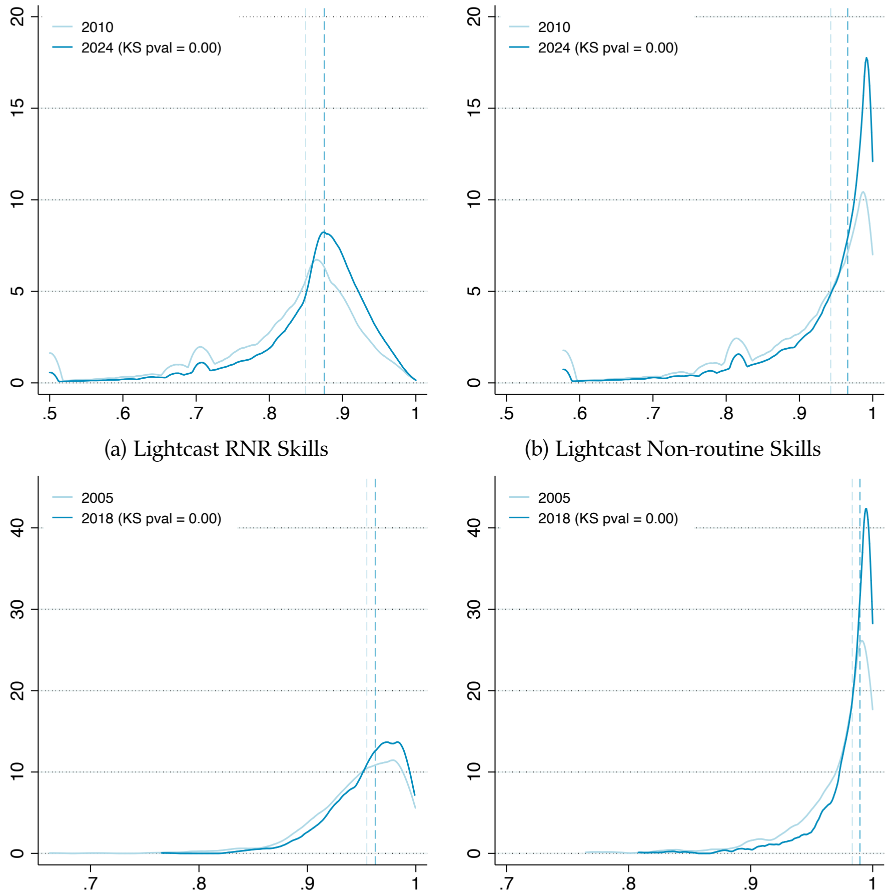
[13] See the Lightcast Open Skills taxonomy documentation (https://lightcast.io/open-skills/categories). In this taxonomy, categories are broad groupings that map roughly to career areas (e.g., Information Technology, Finance, and Health Care).

[14] To simplify the skill dimensions, I use the two non-routine skills (analytical and interpersonal) and combine their routine cognitive and manual skills into one.

[15] I adopt the skill measures from Braxton and Taska (2023) based on the methodology of Hershbein and Kahn (2018). A job posting is classified as requiring analytical skill if its job description includes certain keywords. More specifically, the keywords used to capture analytical skill are: "research", "analy", "decision", "solving", "math", "statistic", and "thinking". The keywords used to capture interpersonal skill are "communication", "teamwork", "collaboration", "negotiation", and "presentation". The keywords used for computer skill are “computer”, or any skill flagged as software by Lightcast.

[16] They also suggest that the evolution of work is not characterized by the active "technification" of jobs through the mixing of these non-routine, general purpose skills.

Figure 4: Density for Skill Mixing Indexes (Cosine Similarities), 2005 vs. 2018



(a) Lightcast RNR Skills

(b) Lightcast Non-routine Skills

(c) O*NET RNR Skills

(d) O*NET Non-routine Skills

Notes: These figures plot the kernel density of different skill mixing indexes over time. The light blue line represents earlier years (2005 for ONET, 2010 for Lightcast), while the dark blue line represents later years (2018 for ONET, 2024 for Lightcast). The Lightcast data are constructed at the job title level, and the O*NET data at the 7-digit occupation level, without employment weighting. The x-axis displays the value of skill mixing indexes with a maximum of 1 by construction. "RNR" indicates routine and non-routine skills that are defined by Acemoglu and Autor (2011). Non-routine skills include non-routine analytical and interpersonal skills, as well as computer skill, as detailed in the online Appendix table A1.

II.C Intensity of Skill Mixing

I confirm the rising skill mixing by constructing a mixing index and show consistency in O*NET data. To do so, I apply an angle-based measure that quantifies the intensity of skill mixing rather than merely the broadening of the skill set. I then demonstrate the consistency of the rise in skill mixing using O*NET data, focusing on the Routine and Non-Routine (RNR) skill dimensions central to the task-based literature.

Data and Measurement: To analyze the intensive margin of skill demand and show robustness, I complement Lightcast with O*NET database, relying specifically on task descriptors derived from job incumbent surveys to ensure longitudinal consistency. To address the challenge of O*NET's staggered updating schedule, I analyze skill changes over broad intervals (2005–2024) during which occupations are updated at least twice. I construct four core Routine/Non-Routine (RNR) skill measures following Acemoglu and Autor 2011) as discussed above, and I normalize these measures using principal component analysis (PCA) and linearly rescaled to the positive unit interval $[0, 1]$ to ensure cardinal comparability.

To quantify the intensity of skill mixing beyond total skill magnitude, I construct an angle-based index within this multi-dimensional space. Formally, I define the degree of skill mixing for an occupation j as the cosine similarity between its skill intensity vector $\mathbf{y}^j$ and a reference norm vector $\hat{\mathbf{v}}$ where all skill requirements are equal:

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

This index captures the angular proximity of an occupation's demand for different skills to a perfectly balanced norm; as the relative importance of distinct skills converges, the index increases toward its maximum value of unity. As such, this index naturally accommodates high-dimensional spaces, remains invariant to proportional shifts in overall skill intensity (thereby isolating compositional changes from pure upskilling), and provides a normalized measure bounded between 0 and 1.

Broad Pattern: Figure 4 shows the density and median values of two skill mixing indexes for 2005 and 2018 using 7-digit O*NET data. The first index covers four routine and non-routine (RNR) skills, while the second focuses on three non-routine skills (analytical, computer, and interpersonal). Panel A shows a modest rightward shift in the RNR skill mixing index, with the Kolmogorov-Smirnov test confirming this shift as statistically significant at the 1 percent level. Panel B reveals a more pronounced rightward shift in the non-routine skill mixing index, with a higher peak in 2018 compared to 2005, indicating a greater growth of non-routine skill mixing in occupations.

The growth in skill mixing is not unique to the choice of non-routine skills and becomes more pronounced when considering employment shares. Online Appendix Figure A2 panel A shows that the rightward shift in the mixing index persists with additional non-routine skills (leadership and design). Panel B combines O*NET data with employment weights from the Occupational Employment and Wage Statistics (OEWS), showing a more pronounced rightward shift in skill mixing indexes when weighted by employment.[17]

III The Role of Breakthrough GPTs

What drives these documented increases in skill mixing? In this section, I examine the impact of breakthrough General Purpose Technologies (GPTs). I focus on breakthrough GPTs due to their defining characteristics–innovations that are both 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).

Below, I first describe the construction of a new measure of exposure to breakthrough GPTs that links patent text to occupational technology descriptors. I then present event study evidence on the effects of these technologies on skill mixing and labor market outcomes.

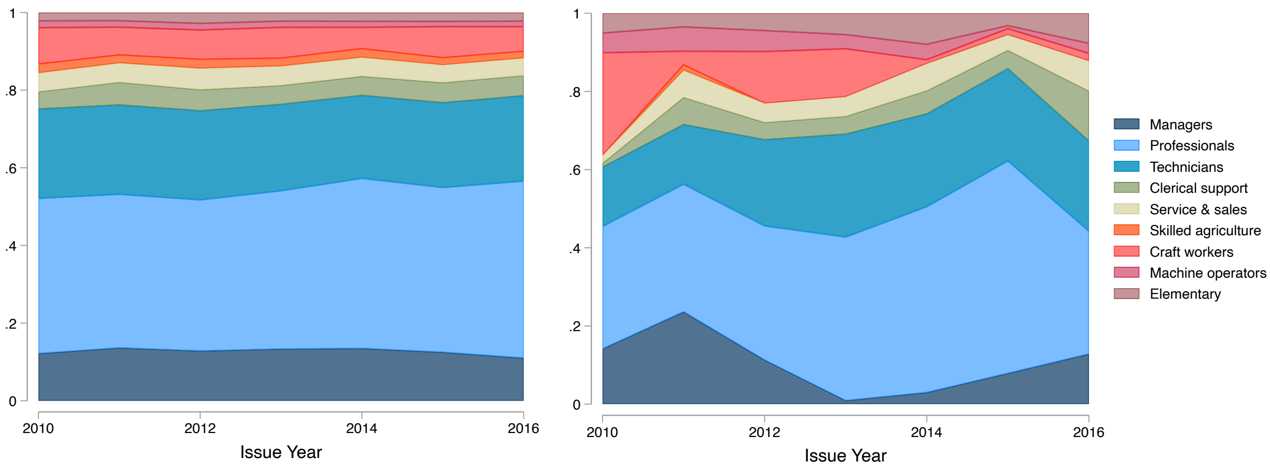
III.A Measuring Exposure to Breakthrough GPTs

To empirically assess the impact of technological shocks, I must first operationalize the concept of a breakthrough GPT. For this purpose, I employ patent data from the USPTO PatentsView database to identify technologies that are both highly innovative and widely applicable. I follow the text-based methodology of Kelly et al. (2021), which uses natural language processing of patent abstracts to identify breakthrough patents based on their textual novelty relative to prior art and their impact on subsequent patents. I then combine this with the citation-based generality index to identify the subset of breakthroughs that are general purpose. Finally, I map these patents to specific occupations using high-dimensional vector embeddings of patent text and O*NET technology descriptors.

First, to capture the "General Purpose" (GPT) dimension, I measure the breadth of a patent's influence across technological fields. I calculate the Generality Index (Gen_p) following the methodology of Hall, Jaffe, and Trajtenberg (2001). For a patent p, generality is defined as $Gen_p = 1 - \sum_c s_{pc}^2$, 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). Intuitively, a higher index indicates that the innovation influences a broader range of subsequent

[17] The OEWS uses 6-digit SOC codes, while O*NET uses 7-digit occupation codes that are based on 6-digit SOC. I match OEWS with O*NET at a 6-digit SOC level and distribute the employment weight evenly for 7-digit O*NET occupations within a 6-digit occupation.

Figure 5: Composition of Breakthrough GPT Patents



(a) Breakthrough Patents (b) Breakthrough GPT Patents

Notes: This figure summarizes the composition of patents classified as Breakthrough GPTs, defined as patents that are: (i) High Generality ($HighGen_p = 1$) if top 25th percentile of Hall, Jaffe, and Trajtenberg (2001) generality index; and (ii) Breakthrough ($BK_p = 1$) if top 10th percentile of Kelly et al. (2021) text-based semantic score. To map these patents to occupations, I compute high-dimensional text embeddings for each patent abstract and match them to detailed "Technology Skills" descriptors from the O*NET database using cosine similarity. These descriptors are then linked to occupations using the official O*NET crosswalk. This mapping enables the construction of a time-varying occupational exposure measure to Breakthrough GPTs.

technological fields, a key feature of a GPT. I define a patent as having "High Generality" ($HighGen_p = 1$) if its index falls within the top 25th percentile of the distribution.

Second, to capture the "breakthrough" nature of an innovation—distinguishing radical advancements from incremental improvements—I utilize the text-based similarity measures developed by Kelly et al. (2021). This approach relies on Natural Language Processing (NLP) of patent abstracts to determine "forward" and "backward" semantic similarity. A patent is classified as a breakthrough if it is textually distinct from preceding patents (low backward similarity) yet highly similar to subsequent patents (high forward similarity), indicating that it initiated a new technological trajectory. I define a patent as a "Breakthrough" ($BK_p = 1$) if its computed breakthrough score is in the top 10th percentile. I then define a *Breakthrough GPT* as any patent that satisfies both criteria simultaneously:

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

The final challenge is to map patent-level innovations to the labor market. I address this by using O*NET technology descriptors, which document the technologies used in each occupation. Specifically, I first generate high-dimensional vector embeddings for the abstracts of all identified breakthrough GPT patents and for the detailed "Technology Skills" descriptors

Table 2: Examples of Breakthrough GPT Patents

Rank	Patent Title	Date	Generality
	Top: High Generality & Breakthrough		
1	Autonomous data machines and systems	2016/5/3	0.9446
2	Airborne fulfillment center utilizing unmanned aerial vehicles for item delivery	2016/4/5	0.9397
3	Nanowire dispersion compositions and uses thereof	2010/6/29	0.9383
4	Device independent authentication system and method	2013/10/29	0.9378
5	Unmanned aerial vehicle navigation assistance	2015/10/27	0.9369
6	Robotic automated storage and retrieval system mixed pallet build system	2014/11/11	0.9349
7	Mast and integral display mount for a material handling vehicle	2016/3/8	0.9349
8	Hand cleansing formulation	2009/11/3	0.9339
9	Three-dimensional traffic flow presentation	2014/3/18	0.9333
10	Automated warehousing using robotic forklifts	2015/2/24	0.9333
	Bottom: Low Generality & Breakthrough		
1	Intervertebral implant	2011/1/4	0.0238
2	Intervertebral implant	2011/1/25	0.0260
3	Blood flow restoration and thrombus management methods	2011/11/29	0.0278
4	Method of utilizing a driverless surgical stapler	2012/9/25	0.0294
5	Circular surgical stapler with mating anvil and shell assembly	2013/4/23	0.0294
6	Methods of treating hypertriglyceridemia	2012/10/23	0.0328
7	Knife/firing rod connection for surgical instrument	2012/12/11	0.0328
8	Surgical instrument	2012/8/7	0.0333
9	Structure containing wound treatment material	2013/12/31	0.0333
10	Surgical stapler	2013/6/4	0.0345

Notes: This table lists the top 10 and bottom 10 patents ranked by their generality index within the sample of "breakthrough" patents. The top panel includes patents with the highest generality scores, as measured by the Hall, Jaffe, and Trajtenberg 2001 index, among patents classified as breakthroughs, defined as being in the top 10th percentile of the Kelly et al. 2021 text-based semantic score. The bottom panel includes patents that satisfy the breakthrough criterion but have the lowest generality scores.

listed in the O*NET database.[18] I calculate the cosine similarity between these embeddings to identify the specific technologies associated with each patent. Subsequently, I utilize the O*NET crosswalk that links these "Technology Skills" to 7-digit occupations. This procedure yields a time-varying measure of exposure, T_o^*, defined as the year of the first exposure to a breakthrough GPT for occupation o.

Composition of Breakthrough GPT Patents: Table 2 reports the ten patents with the highest generality scores among breakthrough patents, along with a comparison set of breakthrough patents with low generality. The high-generality patents in the top panel span core functions in logistics, data handling, materials, and information systems, rather than a single technological

[18]Unlike Continuous Bag of Words (CBOW), which learns static word vectors by predicting a target word from a local context window in a corpus, I use a transformer-based embedding model that produces context-sensitive representations for full text spans, capturing semantic similarity beyond simple co-occurrence. I compute 3,072-dimensional embeddings using OpenAI's `text-embedding-3-large` model.

field. By contrast, the bottom panel contains domain-specific breakthroughs that focus on specialized surgical devices, such as intervertebral implants and stapling instruments. This contrast is informative about the mechanism relating skill mixing to breakthrough GPTs: high-generality breakthroughs create broad platforms that enter many downstream industries, requiring workers in traditional occupations to combine existing tasks with new activities that draw on skills from multiple domains.

Further, Figure 5 plots the evolution of the occupational composition of identified breakthrough GPT patents from 2010 to 2016 and contrasts it with the broader set of breakthrough patents from Kelly et al. (2021). For both groups, patenting activity is concentrated in high-skill occupations: Professionals and Technicians account for the largest shares, Managers contribute a sizable and stable fraction, and clerical, service, and manual groups (craft workers and machine operators) remain relatively small. This pattern is consistent with the view in Kelly et al. (2021) and Autor et al. (2024) that breakthroughs are seed innovations that generate follow-on activities, and it aligns with Figure 2, where high-skill occupations drive broader demand through the diffusion of non-routine skills. However, panel (b) shows that breakthrough GPT patents are broader and more low-skill oriented: relative to panel (a), larger shares occur in clerical support, service and sales, and manual occupations, and the composition varies more over time. This divergence implies that restricting breakthroughs to those that are also GPTs selects innovations whose applications extend beyond core technical high-skill occupations.

III.B The Impact on Skill Mixing

Empirical Strategy: To estimate the causal effect of breakthrough GPT exposure on skill mixing, I employ a dynamic difference-in-differences strategy. Since breakthrough GPTs are not randomly allocated, the obvious identification concern is that the timing of technological adoption might be correlated with pre-existing trends in occupational skill demands. For instance, BKGPTs might specifically target occupations that are already undergoing rapid skill mixing. To address this, I employ an event study design that allows me to test for pre-existing trends. My baseline specification takes the following form:

$$Y_{cot} = \sum_{k=-3}^{8} \beta_k \cdot \mathbb{1}\{t - T^*_{co} = k\} + \delta_{co} + \delta_{ct} + \delta_{ot} + \epsilon_{cot} \tag{2}$$

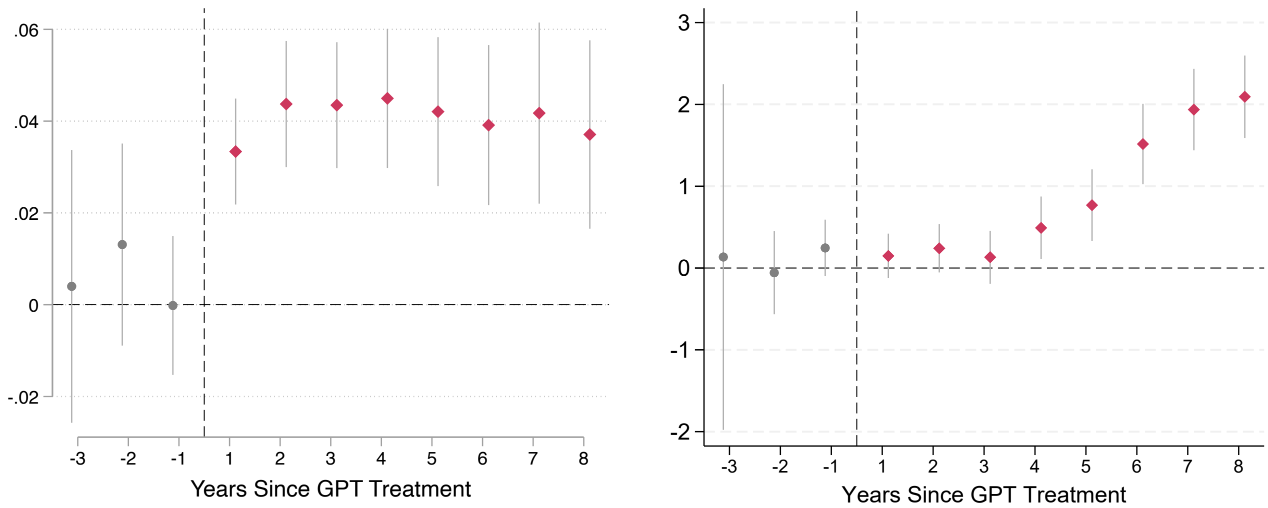
where Y_{cot} is the outcome of interest for the 2-digit occupation group o in the commuting zone c at time t. I use variation at the commuting zone level to study local labor market outcomes, following Acemoglu and Restrepo (2020) and Autor and Dorn (2013). While breakthrough GPT exposure is defined at the occupation level, I assign it to commuting zone $\times$ 2-digit occupation cells to leverage within-zone variation in occupational composition and better adjust for local labor market trends in outcomes.

The specification includes a rich set of fixed effects: commuting zone-by-occupation (δ_{co}) to absorb time-invariant cell characteristics, commuting zone-by-year (δ_{ct}) to control for local economic shocks, and occupation-by-year (δ_{ot}) to control for occupational trends. To address potential biases arising from heterogeneous treatment effects in staggered adoption designs, as well as to accommodate the non-absorbing treatment status of breakthrough GPTs, I estimate the coefficients β_k using the robust estimator proposed by De Chaisemartin and d'Haultfoeuille (2024).

A potential identification concern is that occupations exposed to breakthrough GPTs were already on a different trajectory of upskilling, or that high-exposure commuting zones differ systematically in workforce composition, import competition from China, or the decline of routine jobs; these factors could confound the estimates. In standard long-difference specifications, explicitly controlling for these factors is crucial to isolate the effect of technology. However, the inclusion of high-dimensional fixed effects substantially mitigates this concern. The occupation-by-year fixed effects (δ_{ot}) absorb all common national shocks to specific occupations, such as secular skill-biased technological change or broad changes in labor supply. Similarly, the commuting zone-by-year fixed effects (δ_{ct}) account for all time-varying local shocks, including regional business cycles, local policy changes, local exposure to trade shocks, changes in local industry composition, or shifts in the local supply of skilled labor. Consequently, identification relies only on idiosyncratic variation in exposure within occupations across local labor markets, effectively netting out broader occupational trends and local characteristics that might otherwise drive the results.

Results: Figure 6 presents the event study estimates for the log number of mixed skills required in job postings of exposure to breakthrough GPT patents. Panel (a) displays the average effect. The coefficients for the years leading up to the exposure ($k < 0$) are small and statistically indistinguishable from zero. Upon impact ($k = 1$), there is a sharp and statistically significant increase. Exposure to Breakthrough GPT leads to an immediate increase in the number of mixed skills by approximately 5 percent. Moreover, the effect is not short lived: the estimated coefficient remains similar in size and persists through the remaining periods. Column (1) of Table 3 reports the event-study coefficients in table form. The bottom panel reports the average cumulative (total) effect per treated unit, accumulated over an average of

Figure 6: Dynamic Effects of Breakthrough GPT Exposure on Skill Mixing



(a) Log Number of Mixed Skills

(b) 100 × Skill Mixing Index

Notes: This figure presents dynamic difference-in-differences estimates (β_k) of the effects of exposure to breakthrough GPT on occupational skill mixing. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator to account for staggered treatment and treatment effect heterogeneity. Panels (a) and (b) show impacts on the log number of mixed skills. Panel (a) reports the average effect; Panel (b) reports separate effects by skill level: High-Skill occupations (Managers, Professionals, Technicians, Clerical support) versus Low-Skill occupations (Service and sales, Skilled agriculture, Craft workers, Machine operators, Elementary occupation). Panels (c) and (d) show effects using the Skill Mixing Index (cosine similarity × 100). Panel (c) shows the average effect; Panel (d) presents heterogeneity by occupation skill levels. All specifications include commuting zone-by-occupation, commuting zone-by-Year, and occupation-by-Year fixed effects, and control for lagged demographics, education, and sector composition. The x-axis denotes years relative to first exposure ($k = 0$). Error bars show 95% confidence intervals, clustered at the commuting zone level.

6 post-treatment periods. For the full sample, the average cumulative effect is 0.083 log points, which corresponds to about an 8.7 percent increase in the number of skills mixed.

Columns (2) through (4) of Table 3 address potential specification and measurement concerns. Column (2) imposes a strict consistency requirement by restricting the estimation to the "same switchers" for both the placebo and treatment effects. By enforcing a constant sample for both the treatment and placebo estimates, this specification rules out compositional changes in driving the result.[19] The resulting coefficients are statistically significant and slightly larger than the baseline. Column (3) validates the definition of mixed skills by shifting the base reference year from 2010 to 2011. The resulting coefficients are virtually identical to the baseline in both magnitude and significance, indicating that the findings are not driven by the specific 2010 base year. Finally, column (4) uses the share of job postings requiring

[19]In dynamic difference-in-differences designs estimated via `did_multiplegt_dyn`, the default procedure maximizes the sample size for each time horizon, which can cause the composition of the treatment group to shift across event time.

Table 3: Dynamic Effects of Breakthrough GPT on Skill Mixing

	Log number of mixed skills			Posting Share (%)	Mixing index (×100)		
	(1) All	(2) Same	(3) 2011–24	(4) All	(5) All	(6) Same	(7) 2011–24
Treatment effects							
$k = 1$	0.041***	0.046***	0.017***	0.258**	0.148	0.092	0.917***
	[0.006]	[0.008]	[0.006]	[0.119]	[0.139]	[0.168]	[0.230]
$k = 2$	0.056***	0.068***	0.040***	0.195	0.241	0.114	1.605***
	[0.007]	[0.008]	[0.007]	[0.129]	[0.150]	[0.160]	[0.254]
$k = 3$	0.049***	0.063***	0.043***	0.245*	0.132	0.724***	1.430***
	[0.007]	[0.009]	[0.007]	[0.144]	[0.165]	[0.172]	[0.269]
$k = 4$	0.052***	0.062***	0.049***	0.425**	0.491**	1.171***	1.722***
	[0.008]	[0.009]	[0.008]	[0.171]	[0.195]	[0.214]	[0.252]
$k = 5$	0.046***	0.049***	0.045***	0.119	0.768***	1.335***	1.545***
	[0.008]	[0.009]	[0.009]	[0.179]	[0.223]	[0.250]	[0.253]
$k = 6$	0.043***	0.051***	0.045***	-0.100	1.515***	1.587***	2.042***
	[0.009]	[0.009]	[0.010]	[0.187]	[0.251]	[0.237]	[0.253]
$k = 7$	0.054***	0.059***	0.058***	0.358	1.936***	1.456***	2.636***
	[0.010]	[0.010]	[0.010]	[0.234]	[0.254]	[0.286]	[0.261]
$k = 8$	0.050***	0.054***	0.051***	0.116	2.094***	1.942***	2.865***
	[0.010]	[0.010]	[0.010]	[0.241]	[0.257]	[0.283]	[0.261]
Placebo							
$k = -1$	-0.007	-0.065*	0.021***	-0.193	0.245	0.174	1.516
	[0.008]	[0.037]	[0.008]	[0.177]	[0.176]	[0.228]	[3.023]
$k = -2$	0.020*	-0.018	0.014	-0.311	-0.058	0.089	2.175
	[0.011]	[0.038]	[0.011]	[0.215]	[0.259]	[0.336]	[2.590]
$k = -3$	0.017	0.030	0.002	-0.521*	0.135	-0.442	1.085
	[0.018]	[0.037]	[0.020]	[0.313]	[1.078]	[0.377]	[2.399]
Average total effect							
6 years	0.107***	0.139***	0.092***	0.450*	1.664***	1.952***	5.365***
	[0.013]	[0.017]	[0.014]	[0.254]	[0.331]	[0.334]	[0.585]
czone × occ	✓	✓	✓	✓	✓	✓	✓
czone × year	✓	✓	✓	✓	✓	✓	✓
occ × year	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports event-study difference-in-differences estimates (β_k) of the effects of exposure to a Breakthrough GPT on occupational skill mixing. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator, which accounts for staggered treatment timing and treatment-effect heterogeneity. Columns (1)–(4) report effects on the log number of mixed skills, while columns (5)–(8) report effects using the Skill Mixing Index (cosine similarity ×100). Columns labeled "Low-skill" and "High-skill" report estimates for occupations grouped by skill intensity: High-Skill occupations (Managers, Professionals, Technicians, Clerical support) versus Low-Skill occupations (Service and sales, Skilled agriculture, Craft workers, Machine operators, Elementary occupation). Columns labeled "2011–24" use mixed skills defined over the 2011–2024 period. Event time k is measured in years relative to first exposure ($k = 0$). The table reports the full set of post-treatment effects ($k \geq 1$), placebo leads ($k < 0$), and the average total effect. All specifications include commuting zone-by-occupation, commuting zone-by-year, and occupation-by-year fixed effects, and control for lagged demographic composition, educational attainment, and sectoral shares. Standard errors are clustered at the commuting-zone level and reported in parentheses. Statistical significance is denoted by $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

mixed skills as the outcome variable. This normalization controls the volume of vacancies and shows a similar qualitative pattern.

Further, I examine the impact of breakthrough GPT exposure on the Skill Mixing Index (multiplied by 100), which captures the intensive margin, in Panels (b) and Columns (5)–(7). Following exposure, the mixing index for low-skill occupations rises significantly, reaching an increase of about 2.0 points over the sample period. This confirms that the effect is not merely an addition of skills, but a fundamental rebalancing toward a diversified, mixed set of skill demands. Column (6) imposes the same "same switchers" restriction as in Column (2), which keeps the treatment and placebo samples fixed across horizons; the results remain positive and significant, with an increase of about 1.5 points over the same period. Column (7) shifts the base year for defining mixed skills from 2010 to 2011 and yields an even larger effect (2.7 points), indicating that the rise in the mixing index is not driven by the choice of base year.

Additionally, Online Appendix Table A4 provides further robustness checks, weighting by baseline employment, trimming outliers, and shifting the base year for identifying mixed skill to 2012. The positive relationship between Breakthrough GPT exposure and skill mixing remains qualitatively consistent. Furthermore, Appendix Figure A3 extends the event study over a longer time horizon, which also shows consistent results.

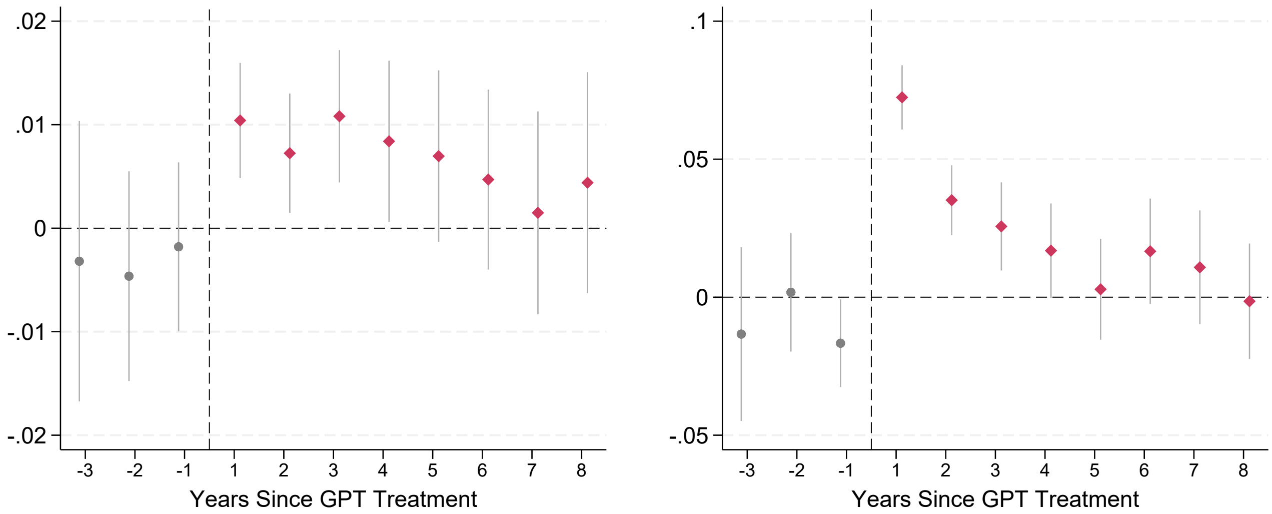
III.C Labor Market Consequences

Does the reorganization of work induced by breakthrough GPTs translate into tangible labor market outcomes? Figure 7 plots the dynamic event study coefficients, while Table 4 reports the corresponding average treatment effects for wages and employment. To construct these measures, I utilize microdata from the American Community Survey (ACS) spanning 2010 to 2023.[20] The sample is restricted to the working-age population (16–64), excluding military, unpaid family, and institutionalized workers, to precisely capture market-driven dynamics.

Panel (a) of Figure 7 and columns (1) through (3) of Table 4 report the dynamic effects on log hourly wages. The baseline estimate in column (1) indicates that exposure to Breakthrough GPT generates a statistically significant wage premium of approximately 1 percent upon impact. This positive effect persists for the first four years before gradually declining and becoming statistically indistinguishable from zero by year seven. The average total effect over the six-year post-treatment period is a statistically significant increase of 1.6 percent.

[20] I utilize IPUMS American Community Survey data (2010–2023). Real hourly wages are calculated as total annual wage income divided by estimated total annual hours (usual weekly hours × imputed weeks), deflated to 2012 dollars using the Personal Consumption Expenditures index. Occupations are harmonized using a consistent census occupation code developed by Autor and Dorn (2013). Individual observations are aggregated to commuting zones using crosswalks from Autor and Dorn (2013).

Figure 7: Dynamic Effects of Breakthrough GPT Exposure on Wages and Employment



(a) Log Hourly Wage

(b) Log Employment

Notes: This figure presents dynamic difference-in-differences estimates (β_k) of the effects of exposure to breakthrough GPT on log hourly wages and employment at the commuting zone $\times$ occupation level. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator to account for staggered treatment and treatment effect heterogeneity. All specifications include commuting zone-by-occupation, commuting zone-by-Year, and occupation-by-Year fixed effects, and control for lagged demographics, education, and sector composition. The x-axis denotes years relative to first exposure ($k = 0$). Error bars show 95% confidence intervals, clustered at the commuting zone level.

Columns (2) and (3) confirm the robustness of this result against compositional biases and an alternative 2011 base year for identifying mixed skills. In both specifications, the estimated wage premiums remain statistically significant and comparable in magnitude to the baseline.

Panel (b) and columns (4) through (6) present the corresponding results for log employment. Like the wage response, employment jumps immediately. Column (4) shows a sharp 7.2 percent increase in employment upon impact. While this initial increase attenuates over time, the effect remains positive and statistically significant for several years, resulting in an average total increase of 5.5 percent over the six-year horizon. As with wages, the robustness checks in Columns (5) and (6)—using the consistent switcher sample and the alternative base year definition—confirm the stability of these large positive employment effects.

Online Appendix Table A5 tests the sensitivity of the labor market results using alternative outcome measures derived from ACS data. The positive wage premium is robust to defining earnings as log weekly wages (column 1) or log annual income (column 2). Similarly, we observe a large and statistically significant increase in the employment rate (column 3), confirming that the expansion in labor demand is evident in the employment-to-population ratio and is not driven solely by population growth.

Table 4: Dynamic Effects of Breakthrough GPT on Wages and Employment

	Log hourly wage			Log employment		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Same	2011–24	All	Same	2011–24
			Treatment effects			
$k = 1$	0.010***	0.010***	0.008**	0.072***	0.072***	0.072***
	[0.003]	[0.003]	[0.004]	[0.006]	[0.007]	[0.006]
$k = 2$	0.007**	0.007**	0.005	0.035***	0.046***	0.035***
	[0.003]	[0.003]	[0.004]	[0.006]	[0.008]	[0.006]
$k = 3$	0.011***	0.011***	0.009**	0.026***	0.027**	0.026***
	[0.003]	[0.003]	[0.004]	[0.008]	[0.011]	[0.008]
$k = 4$	0.008**	0.008**	0.008	0.017*	0.018*	0.017*
	[0.004]	[0.004]	[0.005]	[0.009]	[0.011]	[0.009]
$k = 5$	0.007*	0.007*	0.004	0.003	-0.004	0.003
	[0.004]	[0.004]	[0.005]	[0.009]	[0.011]	[0.009]
$k = 6$	0.005	0.005	0.004	0.017*	0.004	0.017*
	[0.004]	[0.004]	[0.005]	[0.010]	[0.011]	[0.010]
$k = 7$	0.001	0.001	0.001	0.011	0.011	0.011
	[0.005]	[0.005]	[0.005]	[0.011]	[0.011]	[0.011]
$k = 8$	0.004	0.004	0.005	-0.001	-0.003	-0.001
	[0.005]	[0.005]	[0.005]	[0.011]	[0.011]	[0.011]
			Placebo			
$k = -1$	-0.002	-0.002	-0.026	-0.017**	-0.076	-0.017**
	[0.004]	[0.004]	[0.023]	[0.008]	[0.053]	[0.008]
$k = -2$	-0.005	-0.005	-0.005	0.002	-0.110*	0.002
	[0.005]	[0.005]	[0.022]	[0.011]	[0.060]	[0.011]
$k = -3$	-0.003	-0.003	-0.027	-0.013	-0.036	-0.013
	[0.007]	[0.007]	[0.020]	[0.016]	[0.058]	[0.016]
			Average total effect			
6 years	0.016**	0.016**	0.014	0.055***	0.053***	0.055***
	[0.006]	[0.006]	[0.008]	[0.014]	[0.020]	[0.014]
czone × occ	✓	✓	✓	✓	✓	✓
czone × year	✓	✓	✓	✓	✓	✓
occ × year	✓	✓	✓	✓	✓	✓

Notes: This table reports event-study difference-in-differences estimates (β_k) of the effects of exposure to a Breakthrough GPT on occupational wages and employment. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator, which accounts for staggered treatment timing and treatment-effect heterogeneity. Columns (1)–(4) report effects on the log hourly wage, while columns (5)–(8) report effects using log employment. Columns labeled "Low-skill" and "High-skill" report estimates for occupations grouped by skill intensity: High-Skill occupations (Managers, Professionals, Technicians, Clerical support) versus Low-Skill occupations (Service and sales, Skilled agriculture, Craft workers, Machine operators, Elementary occupation). Columns labeled "2011–24" use mixed skills defined over the 2011–2024 period. Event time k is measured in years relative to first exposure ($k = 0$). The table reports the full set of post-treatment effects ($k \geq 1$), placebo leads ($k < 0$), and the average total effect. All specifications include commuting zone-by-occupation, commuting zone-by-year, and occupation-by-year fixed effects. Standard errors are clustered at the commuting-zone level and reported in parentheses. Statistical significance is denoted by $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

Table 5: Average Total Effects of Breakthrough GPT by Occupation Skill Levels

	Skill Mixing		Labor Market Outcomes	
	Log number of mixed skills	Mixing index (×100)	Log hourly wage	Log employment
	(1)	(2)	(3)	(4)
Low-skill occupations				
Average total effect (6 years)	0.140***	1.757***	0.027***	0.085***
	[0.018]	[0.465]	[0.009]	[0.018]
Joint nullity of placebos (p-value)	0.13	0.01	0.42	0.11
High-skill occupations				
Average total effect (6 years)	0.050***	0.317	0.008	0.011
	[0.019]	[0.492]	[0.010]	[0.023]
Joint nullity of placebos (p-value)	0.08	0.00	0.32	0.75
czone × occ	✓	✓	✓	✓
czone × year	✓	✓	✓	✓
occ × year	✓	✓	✓	✓

Notes: This table reports the average total effects over 6 years post-exposure from event-study difference-in-differences estimates of Breakthrough GPT exposure, separately for low-skill and high-skill occupations. Low-skill occupations include Service and sales, Skilled agriculture, Craft workers, Machine operators, and Elementary occupations. High-skill occupations include Managers, Professionals, Technicians, and Clerical support workers. Estimates use the De Chaisemartin and d'Haultfoeuille (2024) estimator accounting for staggered timing and heterogeneous treatment effects. All specifications include commuting zone-by-occupation, commuting zone-by-year, and occupation-by-year fixed effects. Standard errors (clustered at the commuting-zone level) are in brackets. The p-values refer to the joint test that all placebo leads are zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

III.D Heterogeneous Effects by Occupational Skill Levels

The aggregate estimates presented above conceal substantial heterogeneity across occupational levels. As shown in Section II, skill mixing is most prevalent in lower-skill occupations, which also tend to be the primary recipients of mixed skills. In addition, patent composition (Figure 5) reveals that breakthrough GPT patents are more concentrated in lower-skill occupations compared to general breakthrough patents. Motivated by this evidence, I test whether the causal effects of breakthrough GPT exposure are larger for lower-skill occupations. To do so, I estimate heterogeneous treatment effects by splitting the sample into high-skill and low-skill groups, using the same event study framework as specified in equation (2).

Table 5 columns (1) and (2) present the average total effects of Breakthrough GPT exposure on skill mixing, separately for low- and high-skill occupations. The results confirm a clear divergence in how different parts of the occupational hierarchy respond to GPT shocks. For low-skill occupations, exposure leads to a sizable and statistically significant increase in both the number of mixed skills and the mixing index. Over a six-year period, the log number of mixed skills rises by 0.13, while the mixing index (multiplied by 100) increases by 2.05. In contrast, high-skill occupations experience a much smaller rise in the number of mixed skills (0.06) and a slight decline in the mixing index, which is not statistically significant. This result establishes Breakthrough GPT as a driver of the more extensive skill mixing in low-skill occupations, as documented in the descriptive evidence.

This heterogeneity extends to labor market outcomes, as shown in Columns (3) and (4) of Table 4. Among low-skill occupations, exposure to Breakthrough GPT patents leads to a statistically significant 2.7 percent increase in log hourly wages and an 8.6 percent increase in log employment over six years. In contrast, the effects are smaller and statistically weaker for high-skill occupations. The point estimates imply only a 0.8 percent increase in wages and a 1.1 percent increase in employment, with larger standard errors. This pattern suggests that the gains concentrate in occupations where new skills complement existing manual or routine tasks, rather than in occupations that already rely on complex skill bundles.

III.E Discussion of the Impact of Breakthrough GPT

The results that Breakthrough GPTs disproportionately drive skill mixing in lower-skill occupations reinforce the earlier evidence from Section II that low-skill occupations are the main destinations for skill mixing, especially for IT and computer-related skills.[21] They suggest that GPTs are expanding the range of complex skills required in lower-skill occupations. High-skill occupations may originate new skill requirements, but GPT exposure drives the mixing of required skill bundles mainly in lower-skill occupations. For example, a breakthrough in digital logistics does not mean a software engineer must learn new IT skills. Instead, it means a warehouse worker needs to add digital interface skills to their usual manual work. As a result, the biggest shift in skill mix happens in jobs that were once mostly manual or routine but are now taking on cognitive and technical tasks.

Implications for mechanisms: This positive joint response of skill mixing, wages, and employment directly informs the underlying economic mechanisms, which are formalized in Section IV. The premium in both wages and employment suggests that these technologies are likely productivity-enhancing or cost-saving, increasing the demand for labor rather than merely expanding job requirements without corresponding economic returns. By bundling complementary, higher-value skills into lower-skill occupations, breakthrough GPTs raise the marginal product of low-skill labor rather than rendering it obsolete. Consequently, in

[21]This pattern contrasts with a uniform upskilling narrative: Breakthrough GPTs expand the task mix far more in low-skill jobs than in high-skill ones.

Section IV, I develop a multi-dimensional skill matching model that explicitly incorporates multi-dimensional skill complementarity and occupational costs to rationalize these findings.

Difference from automation: Standard automation narratives predict that technology substitutes for routine labor, leading to job displacement. If breakthrough GPTs functioned purely as labor-substituting automation technologies, one would typically expect a trade-off between employment and wages. The results showing that breakthrough GPTs lead to more mixing, higher wages, and higher employment for low-skill workers suggest that these technologies complement low-skill work, raising both pay and employment as employers expand the set of skill demands.

Taken together, the results suggest that breakthrough GPTs increase, rather than replace, the value of low-skill labor. By raising skill requirements in these jobs toward higher-value tasks, such as IT, computer, and other non-routine activities, these technologies increase workers' marginal product and support both higher pay and employment. This contrasts with the standard automation narrative, which emphasizes job displacement through technological substitution.

IV Interpreting the Evidence

To provide a structural interpretation of the increase in skill mixing related to breakthrough GPTs, I develop a multi-dimensional skill matching model with endogenous skill demand. First, firms and workers operate in a multi-dimensional skill space with non-linear production and cost technologies. Second, firms decide their occupations' skill demands before meeting with workers, incurring a cost payable upon operation as in Acemoglu (1999). I use this model to disentangle the mechanisms driving the rise in skill mixing presented in the empirical analysis and assess their implications for aggregate distributions of wages and employment.

IV.A A Multidimensional Skill Framework

Time is discrete. At each period, there is a unit measure of heterogeneous workers that lives forever. Each worker of type i is characterized by a vector of multi-dimensional skills $\mathbf{x}^i = \{x_1^i, ..., x_k^i, \ldots, x_K^i\} \in S \subset \mathbb{R}^{K+}$, 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,

$\mathbf{y}^j = \{y^j_1, ..., y^j_k, \ldots, y^j_K\} \in S \subset \mathbb{R}^{K+}$, which has the interpretation of a vector of skill requirements or skill importance for each of the worker skills.

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:

$$f(\mathbf{x}^i, \mathbf{y}^j) = \left[\sum_{k=1}^{K} (x^i_k \alpha_k y^j_k)^{\sigma^j}\right]^{\frac{1}{\sigma^j}}, \tag{3}$$

where α_k controls the efficiency between worker skill and job skill requirement for a specific skill k, while σ^j governs the elasticity of substitution among different skills for an occupation j.[22] This production technology extends prior multi-dimensional skill matching models (i.e., Lise and Postel-Vinay 2020; Lindenlaub 2017; Ocampo 2022) by incorporating across-skill complementarity controlled by σ^j, in addition to the multiplicative combination of worker and firm attributes intervened by skill efficiencies. This distinction is crucial for our analysis: it allows us to test whether breakthrough GPTs drive skill mixing by fundamentally altering the interdependence of tasks (σ), rather than merely shifting relative factor efficiencies.[23]

Further, firms actively *design* jobs (Acemoglu, 1999), resulting in endogenous skill demand and varying degrees of skill mixing, which offers a structural lens on the empirical increase in skill mixing. Firms set the occupational skill requirements $\mathbf{y}^j$ each period, essentially altering the production technologies that reflect the quality and optimal skill demands of the occupation. Nonetheless, such a job design incurs a cost $C(\mathbf{y}^j) = \sum_{k=1}^{K} (y^j_k)^{\rho^j}$ that is payable upon producing with a worker.[24] This cost captures non-wage expenses of operating an occupation that rise with skill level and job complexity.[25]

Skill Mixing in the Model: The firm's optimal choice of $\mathbf{y}$ determines the equilibrium degree of skill mixing, which can be captured by the skill mixing index. Specifically:

$$Mix(\mathbf{y}) = \frac{\sum_k (\alpha_k x_k)^c}{\sqrt{K}\sqrt{\sum_k (\alpha_k x_k)^{2c}}}, \tag{4}$$

[22]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.

[23]As such, the model explores both the role of changes in relative input efficiency that is the focus of task-based literature and changes in skill complementarity.

[24]This functional form guarantees that the cost is strictly convex and increases with the skill levels chosen by the firm ($\frac{\partial C(\mathbf{y}^j)}{\partial y^j_k} > 0, \frac{\partial^2 C(\mathbf{y}^j)}{\partial (y^j_k)^2} > 0, \forall k$)

[25]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.

Table 6: Parameter Estimates

Parameter	Description	Period 1	Period 2
σ^{low}	Elasticity parameter of skills in production (low skill)	0.53	0.31
σ^{high}	Elasticity parameter of skills in production (high skill)	0.50	0.29
ρ^{low}	Convexity of occupation operation cost (low skill)	3.39	4.10
ρ^{high}	Convexity of occupation operation cost (high skill)	3.37	4.10

Notes: This table shows the exogenously calibrated as well as internally estimated parameters. The data used for the internal estimation are two periods of pooled NLSY79&97 data for workers with information on their pre-market abilities. Period 1 is from 2005–2006 and period 2 from 2016–2019.

where the coefficient $c = \sigma/(\rho - \sigma)$ encapsulates the interplay between technology and costs. Online Appendix B.1 presents the derivation of the model-implied skill mixing index along with its comparative statics.

This formulation reveals four distinct channels through which equilibrium skill mixing can increase, offering a lens to interpret the impact of breakthrough GPTs: (i) an increase in skill complementarity (a decrease in σ), which incentivizes firms to balance skills; (ii) an increase in the convexity of occupation costs (an increase in ρ), which penalizes specialization in a few skills; (iii) a decrease in the dispersion of skill efficiencies ($Var(\alpha_k)$); and (iv) a decrease in the dispersion of worker skills ($Var(x_k)$). Therefore, the empirical finding that breakthrough GPTs increase skill mixing can be rationalized through changes in the model parameters that govern these four channels. Below, I provide a quantitative illustration of the implied parameter changes.

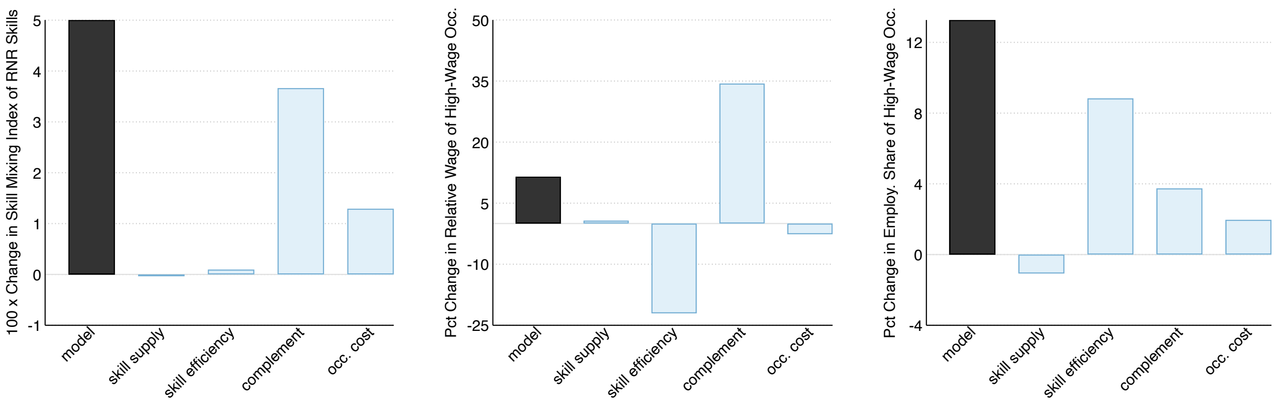
IV.B Interpretation

To understand the mechanisms behind the observed increase in skill mixing, I use the model as an illustrative framework, calibrating it to match both the increase in skill mixing estimated in the event study of breakthrough GPT and key features of the U.S. labor market from 2010 to 2024. I embed this multi-dimensional model within a directed search environment to address the challenge of high-dimensional heterogeneity in matching models.[26] This approach partitions the labor market into submarkets defined by worker skill types, allowing the model to characterize equilibrium tightness and wages independently of the aggregate distribution of workers (Menzio and Shi, 2010), thereby alleviating the need for distributional assumptions.

I utilize the NLSY data to calibrate worker skill endowments across three consolidated

[26] Dimensionality is a challenge inherent in matching models with two-sided heterogeneity. Characterizing equilibrium in multi-dimensional sorting models typically requires specifying the distribution of worker attributes to ensure tractability (e.g., the multivariate normality in Lindenlaub 2017).

Figure 8: Decomposing Channels



(1) Skill mixing (2) Wages (3) Employment

Notes: These figures plot the model-generated changes in skill mixing in low-skill occupations (panel A), changes in the relative wage of high-wage occupations (panel B), and changes in the employment share of high-wage occupations (panel C). Different model channels are shut down in various sequences, and the effect of each channel is calculated by averaging across those sequences. The full model includes all features. The values of skill complementarity in production, cost of skills in occupation operation, efficiency differentials, and vacancy posting costs across the two periods are shown in Table B2. Worker skill supply distribution variation across the periods is calibrated as in Figure **??**.

dimensions: analytical/computer, interpersonal, and routine skills. Standard search model parameters are drawn from existing literature. I fix the skill efficiencies using estimates from Lindenlaub (2017) and Lise and Postel-Vinay (2020), which leaves the parameters governing skill substitution (σ) and occupation costs (ρ) to be determined by data.[27] They are distinguished by exploiting a key model feature: a worker's specific skills drive their output—and thus their wage—but not the cost of occupation design. Therefore, the response of wages to skill differences within the same occupation effectively reveals σ, while ρ is determined jointly to rationalize the intensity of skill mixing. Online Appendix B.3 provides further calibration details.

Magnitudes: Table 6 shows that the calibrated model implies distinct shifts in production technology. Between the two periods, there is a marked increase in the complementarity of skills and a rise in the convexity of occupational costs. As discussed in the theoretical framework, these shifts provide a clear rationale for the observed rise in skill mixing. Figure 8 decomposes the drivers of this increase and shows that it is overwhelmingly explained by higher skill complementarity (73 percent), followed by occupational costs (25 percent).

Finally, I evaluate the aggregate distributional effects of these shifts. The decomposition

[27] As highlighted by Caselli and Coleman (2006), allowing for the endogenous choice of the efficiency of inputs under constraints leaves the elasticity parameters not separately identifiable from efficiencies.

indicates that while complementarity drives the aggregate widening in wage premiums, it is the change in skill efficiencies that primarily drives the increase in employment shares of high-skill occupations (57 percent). This exercise highlights that the technological forces reorganizing job content toward mixed skills are distinct from the efficiency gains that shape aggregate employment distribution.

V 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 where high-skill, non-routine capabilities are increasingly integrated into lower-skill occupations. Leveraging patent and vacancy data from 2005 to 2024, I establish breakthrough GPTs as the key driver of this phenomenon and raise both wages and employment for low-skill workers. Through the lens of a multi-dimensional matching model with endogenous occupation design, increased skill complementarity along with convex occupational costs account for the skill mixing changes and explain a significant portion of the aggregate changes in wage and employment distributions.

These findings carry significant 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, ensuring workers are equipped to reap the returns from the changing demand of 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.

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.

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., 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*.

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.

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. (2021). Technological change and the consequences of job loss. *Working Paper*.

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.

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

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

Davis, S. J., Faberman, R. J., and Haltiwanger, J. C. (2013). The establishment-level behavior of vacancies and hiring. *The Quarterly Journal of Economics*, 128(2):581–622.

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.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, 12(1):210–44.

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.

Juhász, R., Squicciarini, M. P., and Voigtländer, N. (2024). Technology adoption and productivity growth: Evidence from industrialization in france. *Journal of Political Economy*, 132(10):3215–3259.

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.

Menzio, G. and Shi, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4):1453–1494.

Menzio, G. and Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119(3):468–510.

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.

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

In this section, I give more details on data construction for the two primary datasets on job skill demand employed in Section II and III, namely Lightcast (previously known as "Burning Glass") and O*NET (Occupation Information Network. Specifically, I discuss strategies for leveraging the longitudinal information in these datasets with higher precision. I also present an overview of their inherent characteristics, advantages and disadvantages, and how they are cross-walked with other datasets used in the analysis.

Lightcast: Lightcast (formerly "Burning Glass Technologies") is an analytics software company that has developed a comprehensive and detailed dataset derived from online job postings, capturing real-time labor market information, and reflecting the current demand for skills and occupations. One of the key advantages of Lightcast data is its extensive coverage and high-frequency updates. By examining over 40000 online job boards and company websites, it provides a near universe of online posted vacancies; moreover, it provides a level of detail that is rarely matched by other sources of labor market data, such as job titles, employer information, and specific skill requirements. This allows for a very granular analysis of job skill requirements and labor market dynamics across different industries and regions.

The information that Lightcast collected is then parsed and deduplicated into a systematic list of thousands of codified skills. Similar to Hershbein and Kahn (2018) and Braxton and Taska (2023), the dataset that this study uses defines different skills if the codified skills from Lightcast contain relevant keywords. Specifically, the keywords used to capture analytical skill are: "research", "analy", "decision", "solving", "math", "statistic", and "thinking". The keywords used to capture interpersonal skills are "communication", "teamwork", "collaboration", "negotiation", and "presentation". For each occupation, the share of posted vacancies that require a particular skill is then the measure of skill for that occupation, capturing the extensive margin of firm skill demand.

However, like any data source, Lightcast data also has its limitations. For instance, it only covers online job postings, which may not represent the entire labor market, especially for low-skilled jobs or jobs in small firms that do not typically advertise online. It may also have a bias towards certain types of jobs or industries that use online job advertisements more frequently, and online vacancies by nature overrepresent growing firms (Davis, Faberman, and Haltiwanger 2013). One note of Lightcast data is that the measure of skill as introduced above focuses on the extensive margin – whether a job uses a skill or not – this is very different than the level and importance information that O*NET contains.

O*NET: Administered by the U.S. Department of Labor, O*NET is a replacement for the Dictionary of Occupational Titles (DOT). It is more comprehensive and more frequently updated and has been used widely to analyze occupation skill requirements and work settings (i.e., Acemoglu and Autor 2011; Yamaguchi 2012; Autor and Price 2013).

Nonetheless, to use the longitudinal variation from O*NET, the key challenge concerns partial updating – each new version of O*NET only updates an average of 110 targeted occupations among the 970 7-digit occupations. Online Appendix Table A1 lists different versions of O*NET, the release year, and the year composition for 3 of the modules. Specifically, for each release of O*NET, I assign a "Considered Year" such that at least 55% to 60% of occupations are updated after that year.

Moreover, I use 4-year intervals. The last column of online Appendix Table A1 shows the percent of occupations that are updated from the last considered year of data included in the analysis. On average, more than 50 percent of the occupations are updated across the succeeding years included in the analysis.

O*NET contains around 270 descriptors about occupations that are grouped into 9 modules: abilities, knowledge, skills, work context, work activities, experience/education requirement, job interest, work values, and work styles. For my main analysis, I only use descriptors from 3 modules: work context, work activities, and knowledge that are more interpretable as the skill requirements and are consistently evaluated by incumbent workers for each new release. These descriptors come as importance, level, extent, and relevance. To interpret the skill measures as gauging the intensity, I use the importance information, similar to i.e., Acemoglu and Autor (2011) and Guvenen et al. (2020), but the level and importance pieces of information are highly correlated and do not affect the qualitative patterns of skill mixing shown in the paper.

In Section II, I show the longitudinal changes in skill mixing by combining O*NET and ACS datasets. O*NET uses SOC 2000 occupation classification for releases between 2000 and 2010 and SOC 2010 for years after 2010. To link O*NET and ACS, I first bridge SOC codes to the census' OCC 2000 and OCC 2010 codes respectively using crosswalks provided by the Analyst Resource Center and the Bureau of Labor Statistics. Then different years of OCC codes are homogenized using a balanced and consistent panel of occupation codes developed by Autor and Dorn (2013) and updated by Deming (2017). The same code is also used for combining all years of ACS and O*NET data.

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: The table summarizes different versions of the O*NET (Occupational Information Network) database, along with their released year, year division for the 5 modules (work context, work activities, knowledge, skills, abilities), and the considered year for each version. The "Post" and "Before" rows indicate whether the data in each version was collected post or before a particular year. The "Considered Year" column represents the year considered to be corresponding to each release of O*NET based on the year division of data.*

A.2 Details of Skill Measures

In this section, I discuss the choice of skill measures used in the main analysis. Specifically, I show the composition of descriptors of each skill used in the main analysis. I also discuss the composite skill measures' validity and correlation with other measures used in the literature.

Table A2 lists the O*NET descriptors for each of the constructed composite skill measures. The analytical measure corresponds to "non-routine cognitive analytic" and the interpersonal measure corresponds to "non-routine interpersonal" from Acemoglu and Autor (2011). I collapse Acemoglu and Autor (2011)'s "routine cognitive" (the first three items under Routine) and "routine manual" (the last three items under Routine) into a big routine skill, as occupations using these skills have been shown to have had similar labor market dynamics (Autor, Levy, and Murnane 2003; Acemoglu and Autor 2011). I didn't include the "non-routine manual" from Acemoglu and Autor (2011), since it includes descriptors from the "Abilities" module of O*NET that is evaluated solely by job analysts, and for consistency purposes I focus on occupation descriptors that are evaluated incumbents workers.

Further, I include two additional composite skills that are considered to be non-routine. First, I include a "leadership" composite skill that is comprised of descriptors of problem-solving, strategic thinking, teamwork, and communication. They all demand an ability to guide and manage teams, strategize and plan, solve problems, coordinate activities, and communicate effectively within a team or organizational context. Second, I include a "design" composite skill measure centering around technical proficiency and creativity. The composing descriptors entail a strong understanding of design principles, and the ability to draft and layout specifications for technical devices.

Table A2: 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
Design	Design
	Drafting, laying out, and specifying technical devices, parts, and equipment
Leadership	Making decisions and solving problems
	Developing objectives and strategies
	Organizing, planning, and prioritizing work
	Coordinating the work and activities of others
	Developing and building teams
	Guiding, directing, and motivating subordinates
	Provide consultation and advice to others

*Notes: This table shows the detailed O*NET descriptors for skill measures. The Non-routine Analytical and Non-routine Interpersonal skills align with Acemoglu and Autor (2011)'s "non-routine cognitive analytic" and "non-routine interpersonal" skills. A unified Routine skill measure combines Acemoglu and Autor (2011)'s "routine cognitive" and "routine manual" skills, reflecting their similar market trends. The study omits "non-routine manual" to maintain consistency with incumbent worker-evaluated descriptors. Two additional skills, 'leadership' and 'design', are included to capture managerial and creative competencies.*

Table A3: Correlations Among Skill Measures

Skill Measures	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Analytical	1.00								
(2) Routine	-0.45	1.00							
(3) Interpersonal	0.44	-0.49	1.00						
(4) Computer	0.92	-0.27	0.25	1.00					
(5) Math skill	0.50	-0.11	0.12	0.46	1.00				
(6) Social skill	0.34	-0.54	0.61	0.24	0.09	1.00			
(7) Analytical (broader)	0.84	-0.59	0.55	0.68	0.63	0.57	1.00		
(8) Mechanical (broader)	-0.43	0.58	-0.24	-0.38	-0.11	-0.38	-0.49	1.00	
(9) Interpersonal (broader)	0.10	-0.35	0.73	0.02	-0.09	0.70	0.28	-0.22	1.00

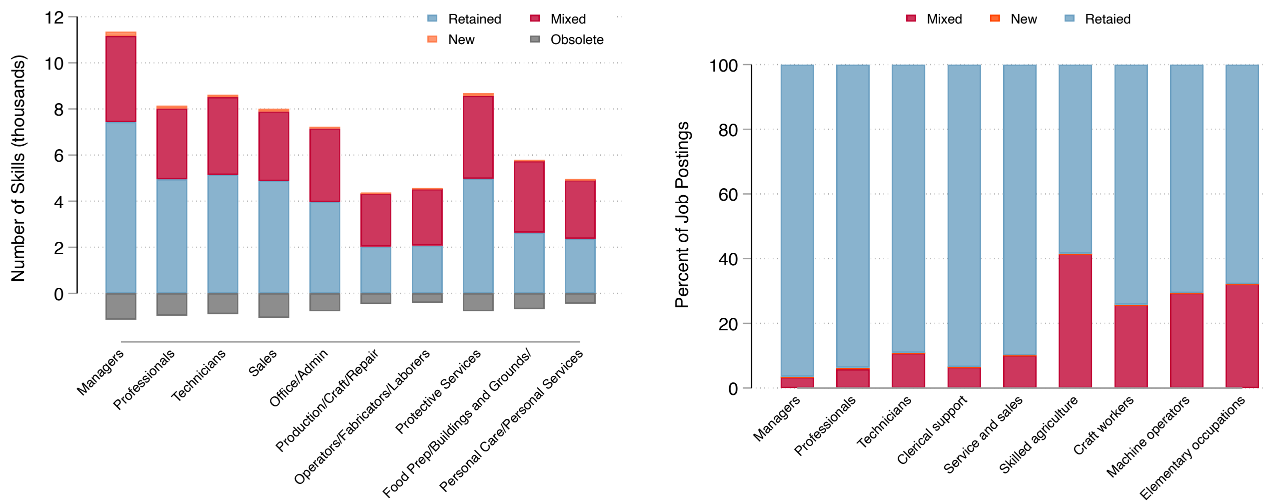
*Notes: This table reports the correlation among different skill measures constructed using O*NET data from 2000-2020. The first four skills measures in rows (1) to (4) are the ones used in the main text and are constructed using the O*NET descriptors shown in Table A1. The next two measures in rows (5) to (6), math skill and social skill are constructed based on Deming (2017). Math skill is the average of 1) mathematical reasoning ability, 2) mathematics knowledge, and 3) mathematics skill. Social skill consists of the average of four variables, 1) social perceptiveness, 2) coordination, 3) persuasion, and 4) negotiation. Rows (7) to (9) contain the broader analytical, mechanical, and interpersonal skills constructed using factor analysis as discussed in online Appendix* **??** *with their specific component variables.*

Table A3 shows the correlation among the chosen skills used in the main analysis, as well as math skill and social skill, which are constructed based on Deming (2017), and broader skill measures skills constructed using factor analysis as discussed in online Appendix **??**. It reveals the analytical skill (row 1), exhibits a strong positive correlation with computer skills (0.92) and a moderate correlation with math skills (0.50). This pattern suggests that positions requiring analytical skills frequently necessitate computer and mathematical proficiency. Interpersonal skills (row 3) indicate a moderate-to-strong positive correlation with social skills (0.61) and broader interpersonal skills (0.73). This correlation suggests that occupations demanding interpersonal skill also emphasize social abilities. These results validate the interpretation of the analytical and interpersonal skills with a strong positive correlation with math and social skills used in other studies.

On the other, a strong negative correlation exists between routine and interpersonal skills (-0.49) and between routine and interpersonal skills (-0.45), indicating that these skill sets rarely overlap in job requirements. The broader skill categories (rows 7 to 9) align well with their narrower counterparts, reinforcing the validity of these categorizations. In sum, there exist specific, identifiable skills in the labor market, some of which are more aligned with each other, but they tend not to overlap, reflecting distinct competencies.

A.3 Alternative Examination of Skill Compositions

Figure A1: Composition of Skills in US Job Postings at 3-Digit Occupations, 2010 to 2024

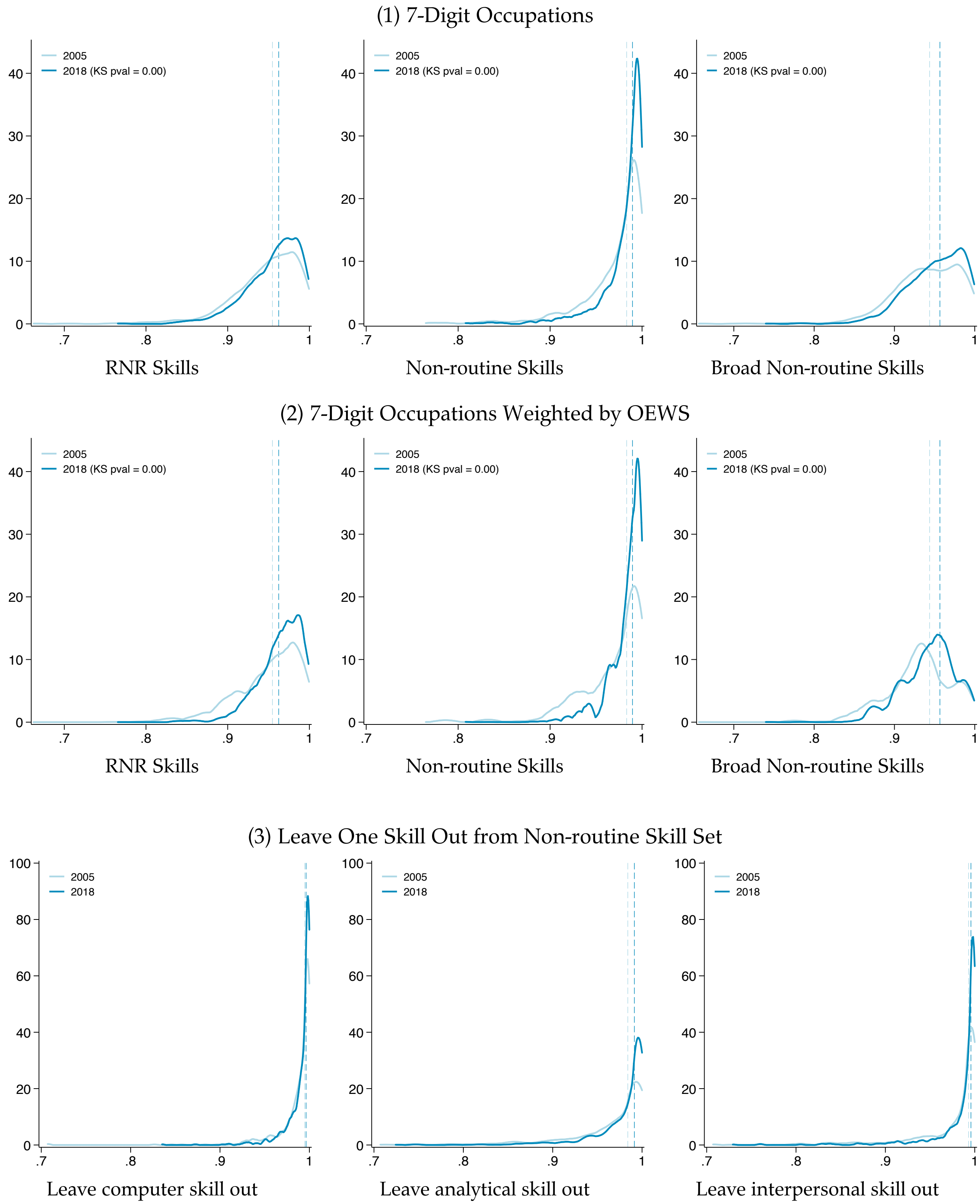


(a) Composition of Skills at 3 Digit SOC

(b) Composition of Posted Skills

Notes: Panel (a) presents the distribution of skill types across major occupational groups in the United States, distinguishing four categories of skills based on their emergence and persistence over time. The occupation classification follows Census OCC codes, harmonized using a balanced and consistent panel of occupation codes developed by Autor and Dorn (2013) and updated by Deming (2017). with calculations first conducted at the 3-digit level and then averaged at the 1-digit level. Values are expressed in thousands, and declining skills are plotted below the horizontal axis for visual clarity. Panel (b) presents the share of job postings in each 1-digit ISCO group that require at least one mixed skill.

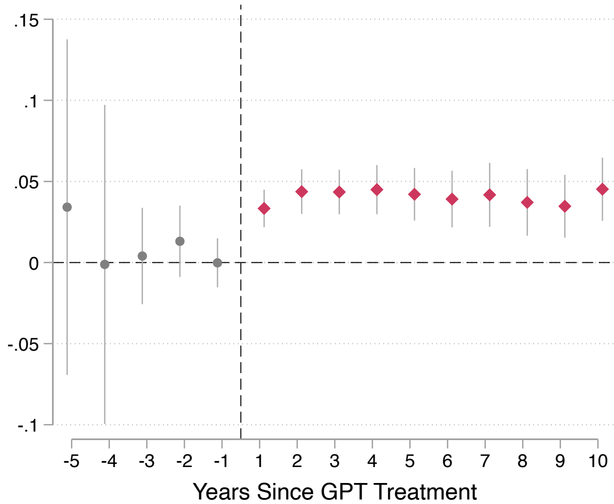
Figure A2: Density for Skill Mixing Indexes (Weighted Cosine Distances), 2005 vs. 2018



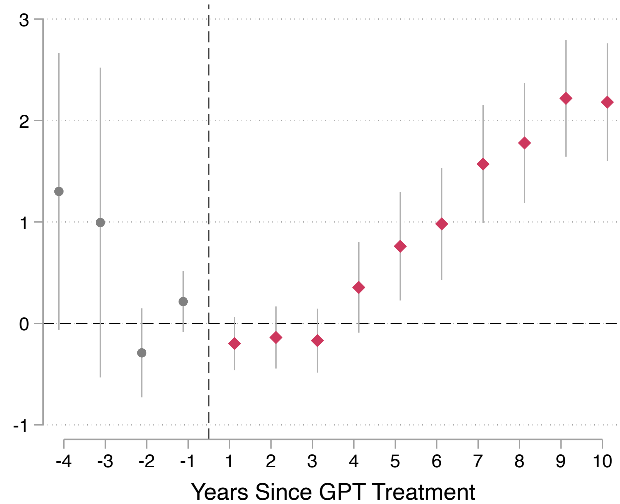
*Notes: These figures plot the PDF of different mixing indexes in 2005 (light blue line) and 2018 (dark blue line). The x-axis displays the value of mixing indexes with a maximum of 1 by construction. These plots are created using O*NET and ACS data merged with occupation codes constructed by Autor and Price (2013) and developed by Deming (2017).*

A.4 Additional Robustness on Event Study Results

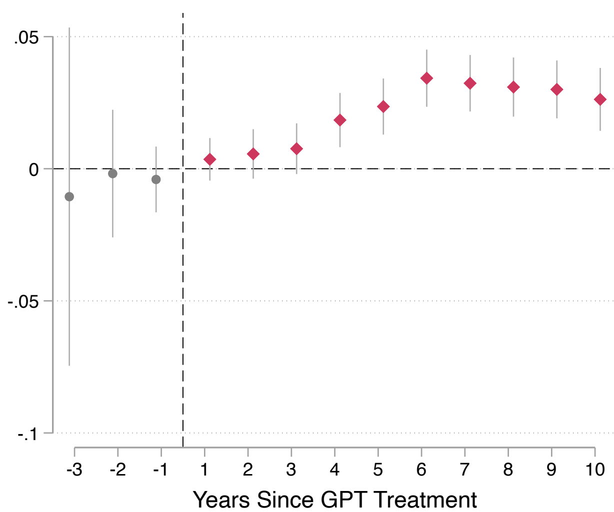
Figure A3: Robustness of the Event Study Results to Longer Time Periods



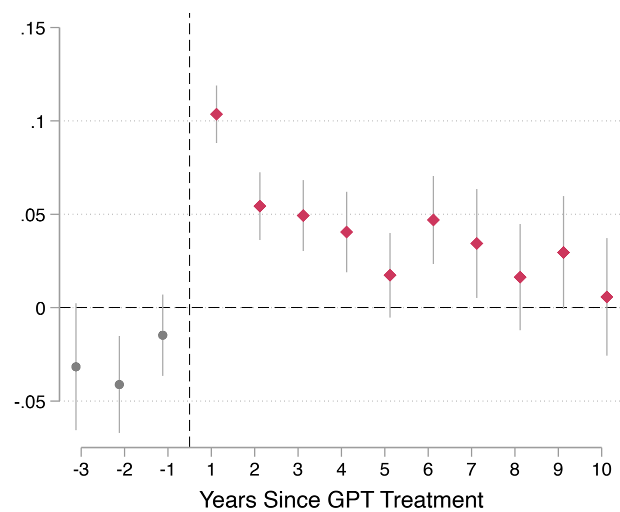
(a) Log Number of Mixed Skills



(b) 100 × Skill Mixing Index



(c) Log Hourly Wage



(d) Log Employment

Notes: This figure presents dynamic difference-in-differences estimates (β_k) of the effects of exposure to breakthrough GPT on occupational skill mixing, wages, and employment. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator to account for staggered treatment and treatment effect heterogeneity. Panels (a) and (b) show impacts on the log number of mixed skills. Panel (a) reports the average effect on log number of skill mixed. Panels (b) presents effects using the Skill Mixing Index (cosine similarity × 100). Panels (c) and (d) presents effects on log hourly wages and employment. All specifications include commuting zone-by-occupation, commuting zone-by-Year, and occupation-by-Year fixed effects, and control for lagged demographics, education, and sector composition. The x-axis denotes years relative to first exposure ($k = 0$). Error bars show 95% confidence intervals, clustered at the commuting zone level.

Table A4: Robustness of Breakthrough GPT Exposure's Impact on Skill Mixing

	Posting Share	Log number of skills mixed			Mixing index (×100)		
Event time	(1)	(2) Weighted	(3) Trimmed	(4) 2012-24	(5) Weighted	(6) Trimmed	(7) 2012-24
				Treatment Effects			
$k = 1$	0.312***	-0.012	0.032***	0.007	0.021	-0.162	0.160
	(0.118)	(0.015)	(0.005)	(0.005)	(0.168)	(0.120)	(0.146)
$k = 2$	0.244*	-0.017	0.038***	0.006	0.326	-0.043	-0.024
	(0.138)	(0.014)	(0.006)	(0.006)	(0.199)	(0.138)	(0.170)
$k = 3$	0.349**	-0.044***	0.049***	0.024***	0.301	-0.333**	0.312*
	(0.140)	(0.015)	(0.006)	(0.007)	(0.218)	(0.146)	(0.172)
$k = 4$	0.407**	-0.068***	0.050***	0.027***	1.273***	0.206	0.768***
	(0.167)	(0.016)	(0.007)	(0.008)	(0.313)	(0.195)	(0.242)
$k = 5$	0.195	-0.113***	0.051***	0.031***	1.806***	0.484**	0.588**
	(0.180)	(0.023)	(0.008)	(0.009)	(0.357)	(0.223)	(0.242)
$k = 6$	-0.043	-0.135***	0.044***	0.031***	1.583***	0.790***	-0.093
	(0.186)	(0.023)	(0.008)	(0.009)	(0.275)	(0.250)	(0.237)
$k = 7$	0.470*	-0.006	0.046***	0.046***	2.708***	1.060***	0.083
	(0.254)	(0.017)	(0.010)	(0.010)	(0.374)	(0.258)	(0.285)
$k = 8$	-0.021	-0.026*	0.047***	0.045***	3.083***	1.190***	0.594**
	(0.252)	(0.015)	(0.010)	(0.010)	(0.377)	(0.256)	(0.293)
				Placebo			
$k = -1$	-0.256*	-0.022**	0.008	-0.008	0.043	0.130	0.150
	(0.150)	(0.011)	(0.007)	(0.007)	(0.247)	(0.130)	(0.264)
$k = -2$	-0.145	-0.008	0.008	-0.002	-0.342	-0.309	-0.530*
	(0.197)	(0.015)	(0.010)	(0.010)	(0.246)	(0.210)	(0.272)
$k = -3$	-0.540**	0.174***	-0.003	-0.011	0.497	1.039	-0.212
	(0.241)	(0.028)	(0.015)	(0.016)	(0.967)	(0.757)	(0.305)
				Average Total Effect			
6 years	0.506**	-0.102***	0.089***	0.050***	2.303***	0.495*	0.570*
	(0.233)	(0.030)	(0.011)	(0.012)	(0.445)	(0.300)	(0.294)
czone × occ	✓	✓	✓	✓	✓	✓	✓
czone × year	✓	✓	✓	✓	✓	✓	✓
occ × year	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports event-study difference-in-differences estimates (β_k) of the effects of exposure to a Breakthrough GPT on occupational skill mixing. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator, which accounts for staggered treatment timing and treatment-effect heterogeneity. Columns labeled "2012–24" use mixed skills defined over the 2012–2024 period. Event time k is measured in years relative to first exposure ($k = 0$). The table reports the full set of post-treatment effects ($k \geq 1$), placebo leads ($k < 0$), and the average total effect. All specifications include commuting zone-by-occupation, commuting zone-by-year, and occupation-by-year fixed effects, and control for lagged demographic composition, educational attainment, and sectoral shares. Standard errors are clustered at the commuting-zone level and reported in parentheses. Statistical significance is denoted by $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

Table A5: Robustness of Breakthrough GPT Exposure's Impact on Wages and Employment

	Log weekly wage (1)	Log income wage (2)	Employment rate (3)
	Treatment effects		
$k = 1$	0.011***	0.030***	0.130***
	[0.003]	[0.005]	[0.014]
$k = 2$	0.006	0.008	0.107***
	[0.003]	[0.006]	[0.016]
$k = 3$	0.011***	0.014**	0.139***
	[0.004]	[0.007]	[0.017]
$k = 4$	0.009**	0.005	0.143***
	[0.005]	[0.008]	[0.020]
$k = 5$	0.005	-0.005	0.114***
	[0.005]	[0.008]	[0.021]
$k = 6$	0.005	0.004	0.140***
	[0.005]	[0.009]	[0.023]
$k = 7$	-0.003	-0.010	0.114***
	[0.006]	[0.010]	[0.023]
$k = 8$	0.002	-0.017*	0.077***
	[0.006]	[0.010]	[0.024]
	Placebo		
$k = -1$	-0.003	-0.014*	-0.080***
	[0.005]	[0.007]	[0.017]
$k = -2$	-0.004	-0.004	-0.052**
	[0.006]	[0.010]	[0.022]
$k = -3$	-0.010	-0.018	-0.119**
	[0.008]	[0.013]	[0.047]
	Average total effect		
6 years	0.014*	0.012	0.268***
	[0.007]	[0.013]	[0.031]
czone $\times$ occ	✓	✓	✓
czone $\times$ year	✓	✓	✓
occ $\times$ year	✓	✓	✓

Notes: This table reports event-study difference-in-differences estimates (β_k) of the effects of exposure to a Breakthrough GPT on wages and employment. Estimates are obtained using the De Chaisemartin and d'Haultfoeuille (2024) estimator, which accounts for staggered treatment timing and treatment-effect heterogeneity. Columns labeled "2012–24" use mixed skills defined over the 2012–2024 period. Event time k is measured in years relative to first exposure ($k = 0$). The table reports the full set of post-treatment effects ($k \geq 1$), placebo leads ($k < 0$), and the average total effect. All specifications include commuting zone-by-occupation, commuting zone-by-year, and occupation-by-year fixed effects. Standard errors are clustered at the commuting-zone level and reported in parentheses. Statistical significance is denoted by $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

B THOERY AND QUANTITATIVE

B.1 Propositions and Proofs

Proposition 1 (Equilibrium Skill Mixing Index)**.** *Consider an occupation characterized by a vector of task requirements* $\mathbf{y} \in \mathbb{R}^{K+}$ *and a worker with skill vector* $\mathbf{x} \in \mathbb{R}^{K+}$*. Given the production technology described in equation (3) with elasticity of substitution* $\sigma \in (0, 1)$ *and a convex cost function* $C(\mathbf{y}) = \sum_k \frac{1}{\rho} y_k^\rho$ *where* $\rho > \sigma$*, the equilibrium degree of skill mixing, measured by the cosine similarity between the optimal task vector* $\mathbf{y}^*$ *and a perfectly balanced reference vector, is given by:*

$$Mix(\mathbf{y}^*) = \frac{\sum_{k=1}^{K} (\alpha_k x_k)^c}{\sqrt{K}\sqrt{\sum_{k=1}^{K} (\alpha_k x_k)^{2c}}}, \tag{5}$$

where $c = \frac{\sigma}{\rho - \sigma} > 0$ *represents the elasticity of task intensity with respect to skill efficiency.*

Proposition 2 (Determinants of Skill Mixing)**.** *Under the conditions established in Proposition 1, the equilibrium degree of skill mixing within an occupation increases if:*

(i) ***Skill complementarity increases:*** *The elasticity of substitution* σ *decreases (implying higher complementarity), provided* σ *remains positive.*

(ii) ***Cost convexity increases:*** *The curvature of the cost function* ρ *increases.*

(iii) ***Skill dispersion decreases:*** *The dispersion of effective skills decreases, specifically if the ratio of effective skills* $(\alpha_k x_k)/(\alpha_h x_h)$ *approaches unity for any task pair* k, h*.*

Proof of Propositions 1 and 2:

Derivation of the Index (Proposition 1). The firm maximizes profit $\Pi = \omega q(\mathbf{x}, \mathbf{y}) - C(\mathbf{y})$. The first-order condition with respect to task intensity y_k is:

$$\omega \frac{\partial q}{\partial y_k} = \frac{\partial C}{\partial y_k} \implies \omega q^{1-\sigma} \alpha_k^\sigma x_k^\sigma y_k^{\sigma-1} = y_k^{\rho-1}.$$

Rearranging terms yields the optimal task intensity $y_k \propto (\alpha_k x_k)^{\frac{\sigma}{\rho-\sigma}}$. Defining the sensitivity parameter $c \equiv \frac{\sigma}{\rho-\sigma}$, we write the optimal vector element as $y_k = \Lambda(\alpha_k x_k)^c$ for some common factor Λ. The skill mixing index is defined as the cosine similarity between $\mathbf{y}$ and the unit vector $\mathbf{u} = [1, 1, \ldots, 1]'$. Substituting the optimal y_k:

$$Mix(\mathbf{y}) = \frac{\mathbf{y} \cdot \mathbf{u}}{\|\mathbf{y}\| \|\mathbf{u}\|} = \frac{\sum_k \Lambda(\alpha_k x_k)^c \cdot 1}{\sqrt{\sum_k [\Lambda(\alpha_k x_k)^c]^2} \cdot \sqrt{K}} = \frac{\sum_k (\alpha_k x_k)^c}{\sqrt{K}\sqrt{\sum_k (\alpha_k x_k)^{2c}}}.$$

This confirms equation (5). $\square$

Proof of Comparative Statics (Proposition 2). To prove conditions (i) and (ii), we examine the behavior of $Mix(\mathbf{y})$ with respect to c. Note that $\frac{\partial c}{\partial \sigma} = \frac{\rho}{(\rho-\sigma)^2} > 0$ and $\frac{\partial c}{\partial \rho} = \frac{-\sigma}{(\rho-\sigma)^2} < 0$. Thus, proving that $Mix(\mathbf{y})$ is decreasing in c is sufficient to prove that mixing increases with lower σ (higher complementarity) and higher ρ (higher cost convexity).

Let $z_k = \alpha_k x_k$. We define $S_1(c) = \sum_k z_k^c$ and $S_2(c) = \sum_k z_k^{2c}$. The index is $M(c) = K^{-1/2} S_1 S_2^{-1/2}$. Taking the log derivative with respect to c:

$$\frac{d \ln M}{dc} = \frac{d}{dc}\left(\ln S_1 - \frac{1}{2}\ln S_2\right) = \frac{\sum_k z_k^c \ln z_k}{S_1} - \frac{\sum_k z_k^{2c} \ln z_k}{S_2}.$$

We define two probability distributions, $p_k = \frac{z_k^c}{S_1}$ and $q_k = \frac{z_k^{2c}}{S_2} = \frac{p_k^2}{\sum_j p_j^2}$. The derivative can be rewritten as the difference in expectations:

$$\frac{d \ln M}{dc} = \mathbb{E}_p[\ln z] - \mathbb{E}_q[\ln z].$$

Notice that the weights q_k put consistently higher probability mass on larger values of z_k compared to weights p_k (since $q_k \propto p_k^2$). Specifically, because p_k and $\ln z_k$ are positively correlated, the covariance $\mathrm{Cov}_p(p_k, \ln z_k)$ is positive. It follows that $\mathbb{E}_q[\ln z] > \mathbb{E}_p[\ln z]$, implying $\frac{d \ln M}{dc} < 0$. Thus, a decrease in σ (lowering c) or an increase in ρ (lowering c) increases skill mixing.

To prove condition (iii), let $u_k = \ln(\alpha_k x_k)$. We can express the index as $M = \frac{1}{\sqrt{K}\|\mathbf{q}\|}$, where $\mathbf{q}$ is a normalized probability vector derived from the inputs. By Jensen's inequality, the Euclidean norm $\|\mathbf{q}\|$ is minimized (and thus M is maximized) when the elements of the underlying distribution are uniform. As the ratio $(\alpha_k x_k)/(\alpha_h x_h) \to 1$, the variance of the inputs decreases, the distribution of task intensities becomes more uniform, and the cosine similarity approaches its maximum value of 1. $\square$

B.2 Equilibrium Definition and Block Recursivity

In this section, I define a block-recursive equilibrium (BRE) for the economy following Menzio and Shi (2011). I further show that the equilibrium of the economy is unique and is block-recursive.

Definition 1 (Block-recursive Equilibrium)**.** *Let $\psi \in \Psi$ be the aggregate state of the economy, which is a distribution of agents across employment status $e = U, W$, skill profiles $\mathbf{x}$, occupational skill requirements $\mathbf{y}$, and output shares ω.*

A block-recursive equilibrium for this economy consists of value functions for both unemployed and employed workers $U(\mathbf{x}) : S \to \mathbb{R}$, $W(\mathbf{x}, \mathbf{y}, \omega) : S \times S \times [0,1] \to \mathbb{R}$, and their respective policy functions $y'_U(\mathbf{x}) : S \to S \times [0,1]$, $y'_W(\mathbf{x}, \mathbf{y}, \omega) : S \times S \times [0,1] \to S \times S \times [0,1]$; firms' policy function $J(\mathbf{x}, \mathbf{y}, \omega) : S \times S \times [0,1] \to \mathbb{R}$ and corresponding policy function $y'_J(\mathbf{x}, \mathbf{y}, \omega) : S \times S \times [0,1] \to S \times S \times [0,1]$; labor market tightness $\theta(\mathbf{x}, \mathbf{y}, \omega) : S \times S \times [0,1] \to \mathbb{R}_+$; and aggregate state $\psi \in \Psi$ such that:

1. *The worker's value functions $U(\mathbf{x})$ and $W(\mathbf{x}, \mathbf{y}$ satisfy (**??**) for all states $\psi \in \Psi$ and $y'_U(\mathbf{x})$, $y'_W(\mathbf{x}, \mathbf{y}, \omega)$ are the associated policy functions respectively*

2. *Firms' value function $J(\mathbf{x}, \mathbf{y}, \omega)$ satisfy (**??**) for all states $\psi \in \Psi$ and $y'_J(\mathbf{x}, \mathbf{y}, \omega)$ is the associated policy function*

3. *The labor market tightness $\theta(\mathbf{x}, \mathbf{y}, \omega)$ in each submarket $(\mathbf{x}, \mathbf{y}, \omega)$ for all states $\psi \in \Psi$ is consistent with free-entry condition in equation (**??**)*

From the above definition of block-recursive equilibrium agents' value functions and policy functions, as well as the market tightness are independent of the aggregate state, only requiring that they are consistent with the aggregate state distribution of agents. Such an equilibrium is easier to characterize analytically and solve numerically. Note a key difference between the above definite of BRE and the one defined in Menzio and Shi (2011). In the economy studied in this paper, because I use the model to study the steady-state equilibrium, the value functions, policy functions, and market tightness are entirely independent of the aggregate state. Whereas Menzio and Shi (2011) studies out-of-steady-state dynamics, the value functions, policy functions, and market tightness still depend on the aggregate productivity shocks but are independent of the distribution of agents across employment status and match-specific shocks.

Now, I show that a block-recursive equilibrium exists and is unique.

Proposition 3 (Existence and Uniqueness of BRE)**.** *Under the model specification of linear utility and invertible and constant returns to scale matching function, also assume that the support for worker and occupation skill profiles S has bounded, then: i) all equilibria are block recurve as defined in definition 1; ii) there exists a unique block-recursive equilibrium.*

Proof of Proposition 3:

The proof first establishes the uniqueness of value functions (U, W, J), as well as policy functions and market tightness $(y'_U, \omega'_U, y'_W, \omega'_W, y'_J, \theta)$; then, the proof establishes their independence from the aggregate state.

Uniqueness: I first show that the value functions for workers and firms as defined in equation (**??**) and (**??**) are contractions. Let $\Theta = S \times S \times [0, 1]$, which is bounded based on the assumption that S is bounded. Let $B(\Theta)$ the space of bounded functions $V : \Theta \to \mathbb{R}$ and the operator associated with the worker or firm value functions denoted by $T : B(\Theta) \to B(\Theta)$. It is straightforward to verify that T satisfies monotonicity and discounting properties:

1. (monotonicity) For $V, V' \in B(\Theta)$, $V \leq V'$ implies $T(V) \leq T(V')$
2. (discounting) For $V \in B(\Theta)$ and $\epsilon > 0$, $T(V + \epsilon) =$

The above conditions establish that the operator T associated with either firm or worker values functions is a contraction bunder Blackwell's sufficient conditions. Therefore, the optimal values workers and firms obtain through dynamic optimization problems are unique.

Next, I show that the policy functions and market tightness are also unique. Since the optimal values firms and workers obtain for their dynamic optimization problems (**??**) and (**??**) are unique, the associated policy functions $(y'_U, \omega'_U, y'_W, \omega'_W, y'_J)$ are also unique due to concavity of the production function defined in equation (3) and workers have linear utility over consumption. To show the uniqueness of market tightness, first note that since it is assumed that the matching function is invertible, one may directly obtain market tightness through the market clearing condition (**??**) with $\theta > 0$. The uniqueness of θ then follows from the uniqueness of firms' value function.

Independence of Aggregate State: In the model economy, workers with different skill profiles $\mathbf{x}$ search in their own market, and firms with different skill requirements $\mathbf{y}$ post jobs in these separated markets, therefore, one can establish that the value functions of firms, workers and the market tightness are all independent of the aggregate state ψ. I establish this argument more rigorously through a backward induction argument as in Braxton and Taska (2021). For this purpose, I introduce back time subscript in the notation.

At the terminal period $t = T$, for an employed worker, the continuation value is zero for $T+1$ onward, so the worker's dynamic programming problem does not depend on the aggregate distribution across states, and is equal to the worker's share of output $W_T(\mathbf{x}, \mathbf{y}, \omega) = \omega f(\mathbf{x}, \mathbf{y})$.

Similarly, the firm's value function also remains independent of the aggregate distribution $J_T(\mathbf{x}, \mathbf{y}, \omega) = (1-\omega) f(\mathbf{x}, \mathbf{y})$. As a result, through the free entry condition in equation (**??**), the market tightness $\theta_T(\mathbf{x}, \mathbf{y}, \omega)$ is also independent of the aggregate distribution.

Firms at $T-1$ make occupation design choices $\mathbf{y}$ to solve the firm dynamic programming problem in equation (**??**); workers at $T-1$ make labor market search choices over occupations $\mathbf{y}$ to solve the worker dynamic programming problem in equation (**??**); As long as $\mathbf{y}$ is within a bounded interval, the extreme value theorem assures at least one solution to this problem. This process is repeated stepping back from $t = T-1, ..., 1$, which completes the proof. *Q.E.D.*

B.3 Model Calibration

I use the same combination of NLSY 79 & 97 and ONET data as in Section III to quantify the model with three skills. NLSY data informs us about worker abilities ($\mathbf{x}$) and captures changes in employment and wages, while ONET provides occupational skill requirements ($\mathbf{y}$). For consistency, the sample is restricted to data from 2005–2006 and 2016–2019 and includes only workers with available skill information.[28] Finally, for both worker and job skill profiles, I consider the same set of skills (analytical, computer, interpersonal, routine) as in Sections II and III, only that I combine analytical and computer skills to have a three-dimensionality feasible for quantitative analysis ($K = 3, k = \{\text{analytical/computer } (a), \text{ interpersonal } (p), \text{ routine } (r)\}$).[29]

Considering the potential influence of skill supply on skill mixing, I calibrate worker skill distribution $G(\mathbf{x})$ across two periods to reflect choices of occupations and college majors. Workers accumulate skills at a rate γ_j based on the gap between their current skills and the requirements of their occupation or college major, using NLSY's occupational/major data. The adjustment rate γ_j varies by skill and the direction of adjustment using estimates from Lise and Postel-Vinay (2020).[30]

To map occupations and workers in the model to the data, I set grid points as follows. I classify occupations into high- and low-wage, as in Section III, with the former group including managerial, professional, and white-collar occupations, and the latter blue-collar and service occupations. The grid point for an occupation's requirement of a skill y_j is set such that moving up one grid corresponds to 50 percent of the average observed value of y_j for that occupation.[31] On the worker side, workers are classified based on their skill level x_j: those with skills above the average are deemed high type and assigned the mean of the above-average values; those below the average are considered low type and assigned the mean of the below-average values.[32]

Functional Forms: The functional forms are chosen as follows. The multi-dimensional skill production function is defined in equation (3), which extends the multi-dimensional matching literature (i.e., Lise and Postel-Vinay 2020; Lindenlaub 2017; Ocampo 2022) by

[28]NLSY 1997 was conducted annually during 2005-2006 but biannually in 2016-2019, as was NLSY 1979 for the later period. The sample sizes for these two periods are 30,654 and 43,340 respectively.

[29]As I merge analytical and computer skills into one for calibration using their average values, I denote this combined skill as "analytical/computer".

[30]Details of this calibration are available in online Appendix B.5. Workers' skills adjust downward when unemployed but cannot be lower than their initial endowments. For skill changes while in college, I specify that workers spend on average 3 years learning the skills of their majors.

[31]As the model calibration uses data of two periods with a consistent grid, I determine grid points by averaging the occupation's median values across both periods.

[32]With three chosen skills, there are 8 worker types in the model.

Table B1: Moments and Model Match

	First Period		Second Period	
	Data	Model	Data	Model
Worker moments				
Relative wage of high type				
Analytical/computer	1.46	1.36	1.60	1.63
Interpersonal	1.05	1.17	1.20	1.25
Routine	1.12	1.48	0.92	1.27
Wage return of skill mixing (untargeted)	0.07	0.04	0.07	0.04
Unemployment rate	0.05	0.07	0.04	0.07
Occupation moments				
Relative wage of high skill	1.30	1.27	1.56	1.55
Corr. wage and abilities (low-wage)	0.23	0.28	0.49	0.39
Corr. wage and abilities (high-wage)	0.35	0.27	0.60	0.64
Employment share (low-wage)	0.43	0.32	0.37	0.12
Employment share (high-wage)	0.57	0.68	0.63	0.88
100 $\times$ Skill mixing (low-wage)	97.54	94.51	98.96	99.14
100 $\times$ Skill mixing (high-wage)	95.74	96.44	94.12	94.73

Notes: This table reports the average values of the targeted moments both in the data and through model simulation. The data used for the moment calculation and for SMM estimation are two periods of pooled NLSY79&97 for employed workers: period 1 from 2005–2006 and period 2 from 2016–2019. Two types of moments are included. The worker moments include the relative wage of high type workers as well as the unemployment rate. The occupation moments include the relative wage of high skill occupations, the employment share and the skill mixing index of RNR skills in low and high skill occupations.

incorporating both skill-specific efficiency of matching α_k and cross-skill complementarity σ^j. For the occupation operation cost function, I apply a simple and flexible formulation, $C(\mathbf{y}) = \tau[\sum_{k=1}^{K}(y^k)^\rho]$, which is uniform across all occupations.[33] Here, ρ determines the convexity of the cost function relative to skill levels, and τ sets the cost scale. The matching function adopts a standard Cobb-Douglas format, $M(s, v) = \mu s^\eta v^{1-\eta}$, where η measures the elasticity of matches in relation to total search effort and μ reflects matching efficiency, leading to a job finding rate of $p(\theta) = \mu\theta^{1-\eta}$ and a vacancy filling rate of $q(\theta) = \mu\theta^{-\eta}$.

The calibration of parameters falls into three categories. For parameters that regulate the search environment, I follow closely the conventions of the search literature. For skill adjustment and efficiency parameters, I draw on estimates from the multi-dimensional matching literature. Finally, I internally estimate the production technology parameters, which govern the elasticity of substitution across skills and the scale and convexity of the operational costs of skills, using Simulated Methods of Moments (SMM).

[33] Besides technical convenience, the functional form also implies that for a given cost, firms need to trade off the choice of altering different skill intensities.

Table B2: Parameter Estimates

Parameter	Description	Value	
	Panel A. Externally calibrated - search		
β	Discount Rate	0.96	
δ	Job separation rate	0.10	
ω	Worker share of surplus	0.60	
b	Unemployment benefit as a share of output	0.42	
η	Elasticity of the matching function	0.50	
μ	Matching efficiency	0.65	
	Panel B. Externally calibrated - skill adjustment	Up	Down
γ_a	Annual adjustment speed of analytical/computer skill	0.36	0.10
γ_p	Annual adjustment speed of interpersonal skill	0.05	0.00
γ_r	Annual adjustment speed of routine skill	1.00	0.36
	Panel C. Externally calibrated - skill efficiency	2005	2018
α_a	Skill efficiency of analytical/computer skill	0.63	0.95
α_p	Skill efficiency of interpersonal skill	0.05	0.08
α_r	Skill efficiency of routine skill	0.14	0.06
	Panel D. Internally estimated	2005	2018
σ_{low}	Elasticity parameter of skills in production (low skill)	0.53	0.31
σ_{high}	Elasticity parameter of skills in production (high skill)	0.50	0.29
ϕ_{low}	Convexity of occupation operation cost (low skill)	3.39	4.10
ϕ_{high}	Convexity of occupation operation cost (high skill)	3.37	4.10

Notes: This table shows the exogenously calibrated as well as internally estimated parameters. The data used for the internal estimation are two periods of pooled NLSY79&97 data for workers with information on their pre-market abilities. Period 1 is from 2005–2006 and period 2 from 2016–2019.

External Calibration: The model period is a year. Given that all agents are risk-neutral, the discount rate β is assigned a value of 0.96, corresponding to an annual interest rate of 4 percent. The job separation rate δ is set at 10 percent as in Shimer (2005). For employed workers, their share of output ω is set at 0.6, mirroring the labor share of GDP in 2005. For unemployed workers, the unemployment benefits b is set at 41.5 percent of the earning loss of lowest-paid occupations, following the estimates of Braxton, Herkenhoff, and Phillips (2020). The elasticity of the matching function η is set at 0.5 as is standard, and the matching efficiency μ is set to 0.65, as in Mercan and Schoefer (2020). Table B2 panel A summarizes these externally calibrated parameters.

I calibrate the speed of skill adjustment (γ_j) and the skill efficiencies (α_k) following Lise and Postel-Vinay (2020) and Lindenlaub (2017), as detailed in Table B2 panels B and C. The calibration aligns the adjustment of analytical/computer, interpersonal, and routine skills with the cognitive, interpersonal, and manual skills detailed in Lise and Postel-Vinay

(2020).[34] Analytical/computer skill adjusts upward two times faster than it depreciates, while interpersonal skill changes slowly in both directions. Routine skill adjusts most rapidly in either direction. I linearly interpolate Lindenlaub (2017)'s estimates of skill efficiencies for my period of analysis.[35] Between 2005 and 2018, the productivity of analytical/computer and interpersonal skill in matching worker abilities with job skill requirements increased by about 60 percent. In contrast, the productivity of routine skill saw a decrease of more than 50 percent.

Internal Estimation: For the internal estimation, the SMM procedure entails solving the agents' steady-state policies and simulating a cohort of workers for $T(T > 80)$ periods, resulting in a distribution of labor market outcomes. The parameters are then estimated minimizing the distance between simulated and empirical moments.[36] The estimation targets 11 moments as shown in Table B1 for both periods of data that include: i) the relative wage of the high-type worker for each skill; ii) the unemployment rate; iii) the relative wage of high-skill occupation; iv) the within-occupation correlation between wages and worker abilities; v) the share of employment across occupations; and vi) the skill mixing index of RNR skills of occupations.[37] The model does a decent job of matching all the moments, and replicates the non-targeted wage returns of skill mixing in Section III.

The model parameters are jointly identified from the moments, for which a concise summary of the key information for identification is given below with a more detailed discussion in online Appendix B.4. I first identify the complementarity parameter of skills in production σ targeting the correlation of within-occupation relative wages and worker skills. The cost parameter ρ is then estimated by leveraging the firm's optimization conditions in skill mixing. Conditional on parameters estimated at the production side, the employment distribution and relative wages further aid in estimating τ. Lastly, the unemployment rate disciplines the vacancy posting cost c.

Table B2 panel D presents the internally estimated parameters, which indicate considerable technological shifts between the two periods. For the initial period, the estimated σ is 0.6 for

[34] Lise and Postel-Vinay (2020)'s estimates are presented on a monthly basis, which I have adjusted to an annual scale.

[35] Lindenlaub (2017)'s estimates span from 1990 to 2010.

[36] Online Appendix **??** provides further details on the numerical implementation.

[37] All moments are directly computed from the two periods of data from NLSY, except for unemployment, for which I use the statistics from the Bureau of Economic Analysis (BEA) to avoid the age composition effects present in NLSY. For example, by the late 2010s, a larger segment of the NLSY 79 cohort was above age 50, making them more likely to be out of the labor force. Additionally, the unemployment rate from NLSY, derived from the number of jobs held since the last survey, averages 9 percent, notably higher than BEA data. However, this decision primarily affects vacancy posting cost parameters.

low- and high-wage occupations, suggesting that skills are substitutable in production. In the late 2010s, there was a significant rise in skill complementarity in production, reflected in the reduction of σ to 0.3 for both types of occupations. Firms also encounter rising costs of skills in occupation operation, as reflected in the increase of both the scale and the convexity of the cost function (τ and ρ). As discussed in Section IV.B, this increased complementarity as well as the operational costs of skills intensifies firms' incentives to mix skills. Lastly, the cost of posting vacancies remains relatively constant.

B.4 Identification of Parameters

I begin by estimating the elasticity parameters in production and occupation operation cost, denoted by σ and ρ. As highlighted by Caselli and Coleman (2006), the challenge arises when allowing for the endogenous choice of the efficiency of inputs under constraints, as the elasticity parameters cannot be separately identified. To overcome this challenge, I estimate σ using the *relative wage within occupation* instead of relying on absolute wage levels.

Specifically, based on the model, the wage that workers receive per period is given by the share ω of the output of the worker-firm match, reduced by the occupation design cost, formulated as $w(\mathbf{x}, \mathbf{y}) = \omega f(\mathbf{x}, \mathbf{y}) - C(\mathbf{y})$. Consequently, within each occupation, the difference in wage relative to a base worker type $\Delta w(\mathbf{x}, \mathbf{y})$ can be articulated as follows:

$$\Delta w(\mathbf{x}, \mathbf{y}) = \omega \left[\sum_{k=1}^{K} (x^k y^k)^{\sigma} \right]^{\frac{1}{\sigma}} - A, \tag{6}$$

where A is occupation-specific and does not depend on the cost parameter τ or ρ. This formulation enables the identification of σ independent of the cost parameters. To carry out the estimation equation (6), I first adjust the wage for occupation fixed effects in order to account for A and ω. Next, I compute the within-occupation difference of the adjusted wage relative to the lowest skill type worker.[38] Last, I target the correlation between this adjusted within-occupation relative wage and worker abilities $\mathbf{x}$.[39]

I now turn to the identification of the cost parameters ρ and τ. To begin with, note that the first-order condition of firms' optimization problems in the submarket $(\mathbf{x}, \mathbf{y})$ can be simplified in ratios to $\frac{y_h}{y_k} = \left(\frac{x_h}{x_k}\right)^{\frac{\sigma}{\rho - \sigma}}$, a relationship that exclusively depends on the parameters σ and ρ. With σ already estimated, I then target the skill ratio y_j / y_k, which aligns with the moment of the degree of hybridization of occupations. Further, for employed workers, the distribution of employment across various occupations is governed by wages $w(\mathbf{x}, \mathbf{y})$. Given the parameters described above, this functional relationship allows the estimation of τ.

Lastly, given the calibrated unemployment benefits b, the parameters of the matching, production, and cost functions, equation (**??**) reveals that the probability of exiting unemployment only depends on the vacancy posting cost. By targeting the unemployment level, c is identified.

[38] Refer to Section IV.B for an in-depth discussion on how worker skill types are calibrated.

[39] According to equation (6), σ can be identified from the correlation of any skill with the adjusted wage, which is what I use as the target.

B.5 Calibration of Skill Supply

I carry out the calibration of two key aspects of skill supply variation: the Markov probability of worker skill adjustment in a steady state equilibrium and the variation in worker skill supply spanning two data periods that the model aims to align with two steady states. I will first delve into the details of the skill variation between data periods and then explore the skill evolution within a model period as guided by the Markov process, following the approach of Lise and Postel-Vinay (2020).

Across-period Skill Supply Variation: Considering the potential influence of skill supply variation on skill mixing, I calibrate the model to reflect workers' choices in occupation, college major (if attended), and employment status, in line with the approach of Lise and Postel-Vinay (2020). This calibration introduces variation in worker skill supply across two periods. Worker skills are adjusted based on the requirements of an occupation or a college major; they increase if the requirements exceed the original skills and decrease if the requirements are lower or if the worker is unemployed. The speed of this adjustment is asymmetric and skill-specific.

Specifically, following the estimates from Lise and Postel-Vinay (2020), as presented in online Appendix Table B3, a worker's skills accumulate at a rate of γ_j times the gap between the worker's skill j and the occupation's requirement for that skill each year. The value of γ_j depends on whether it relates to learning or depreciation (upward or downward accumulation). Additionally, workers can lose skills when not employed, with unemployment treated as requiring a zero level for all skills. However, I specify such that a worker's skill level cannot fall below their initial endowments. For changes in skills while in school, I specify that workers spend an average of three years learning the skills of their majors.

I incorporate two modifications into this framework. First, since Lise and Postel-Vinay (2020)'s estimates are based on weekly data, I adjust them by multiplying by the number of working weeks, set at 47. Second, I align Lise and Postel-Vinay (2020)'s estimates of cognitive, interpersonal, and manual skills with my analysis's categories of analytical, interpersonal, and routine skills.[40] Since Lise and Postel-Vinay (2020)'s estimates do not include computer skills, I use their cognitive skill estimates as a proxy.

In calculating the skill adjustment, I first standardize both worker skills and occupation skill requirements. Then, for example, if a worker is employed in an occupation that requires

[40] Their exclusion restriction imposes that (i) the ASVAB mathematics knowledge score only reflects cognitive skills; (2) the ASVAB automotive and shop information score only reflects manual skills; (3) the Rosenberg self-esteem score only reflects interpersonal skills.

a standard deviation higher in analytical skill compared to the worker's analytical skill, the worker will accumulate 0.36 standard deviations of analytical skill in a year due to learning on the job. Conversely, if a worker's interpersonal skill is higher than required, it will decrease by only 3×10^4 standard deviations, almost remaining unchanged, as interpersonal skills are estimated to be very hard to lose.

Table B3: Annual Skill Learning and Depreciation Rate

O*NET Measures	NLSY Measure	$\gamma^{\text{learn}}_{\text{school}}$	γ^{up}_{j}	γ^{down}_{j}
Analytical	AFQT score	0.33	0.36	0.10
Interpersonal	Deming (2017) social skill	0.33	0.05	0.00003
Routine	ASVAB mechanical	0.33	1.00	0.36
Computer	OCC/Major's 2005 value	0.33	0.36	0.10

*Notes: This table illustrates for each O*NET skill measure, its corresponding skill measure using NLSY79&97 data, and the learning and depreciation rate for these different skills. The AFQT is the same as the one used by Altonji, Bharadwaj, and Lange (2012) followed by Deming (2017), which controls for age-at-test, test format, and other idiosyncrasies. Deming (2017)'s social skill measure consists of sociability in childhood and sociability in adulthood in NLSY79, and two questions from the Big 5 inventory gauging the extraversion in NLSY97. The average of workers' ASVAB mechanical orientation and electronics test scores are used for mechanical skill. Since ASVAB scores are not available for the NLSY97 survey, they are imputed based on predictive regression using the NLSY79 survey. Workers' occupations' or college majors' O*NET computer skill scores in the year 2000 are used as their endowed computer skill. The skill accumulation/depreciation rate is directly from Lise and Postel-Vinay (2020)'s estimates based on monthly data converted to annual values. Skill learning/depreciating while attending college is specified to be 33% per year.*