Automation without Borders: Generative AI and Cross-Border Service Outsourcing

Neha V. Betai Maggie X. Chen*

February 5, 2026

Abstract

The outsourcing of routine cognitive work to developing countries has been a resilient component of globalization. This paper investigates how generative artificial intelligence (AI) reshapes cross-border service outsourcing by tracing firm and worker responses to the diffusion of AI in a global market for digital services. Using granular transaction data linking clients, workers, and skills from 2022–2024, together with a novel capability-based measure of AI automation exposure spanning 2,700 skill categories, we document a sharp contraction in cross-border outsourcing in highly exposed skill areas. Outsourcing volume falls by roughly 35 percent, with declines concentrated in developing-country destinations that previously specialized in codifiable work. These aggregate effects mask substantial reorganization. Firms consolidate remaining outsourcing into fewer, larger, and more complex contracts. Exposed workers complete fewer contracts but earn more per engagement, as a subset reallocates toward AI- and interaction-intensive tasks. Adjustment is uneven, leading to a 22-percentage-point increase in within-skill concentration of activity. A task-based trade framework rationalizes these patterns by showing how generative AI compresses cross-border wage and productivity arbitrages in routine tasks, shifting comparative advantage and redistributing gains from trade in digital services.

JEL codes: F1, O3

Keywords: Generative AI, international outsourcing, services, reskilling, concentration

*George Washington University. Neha Betai: nehabetai96@gwu.edu; Maggie Chen (Corresponding author): xchen@gwu.edu. We thank seminar and conference participants at IGC-Yale's 2025 Firms, Trade, and Development Conference, Paris School of Economics "AI, Digitization and Trade Patterns" Workshop, Inter-American Development Bank, World Bank, International Finance Corporation, HKUST-Fudan-NUS Conference on International Economics, and Lingnan-Nottingham Workshop on International Trade for valuable comments and suggestions.

1 Introduction

The rise of generative artificial intelligence (AI) marks a potential turning point in the global organization of work. Over the past several decades, advances in information and communication technologies enabled firms in high-income countries to unbundle production and offshore routine cognitive tasks to lower-wage workers in developing economies. Digitally deliverable services became one of the fastest-growing and most resilient components of international trade, fostering new forms of export-led growth (World Bank, 2023). This "trade in tasks" reshaped labor markets on both sides of the transaction, creating export opportunities for workers in developing economies while enabling cost savings and specialization for firms in developed economies (Grossman and Rossi-Hansberg, 2008; Baldwin, 2019). Generative AI threatens to upend this equilibrium. By automating precisely the routine, codifiable tasks that underpinned developing countries' comparative advantage, the technology may compress the cross-border wage arbitrage that sustained global digital outsourcing, altering not only how work is performed, but where it is produced.

This paper investigates how generative AI reshapes cross-border service outsourcing, tracing the effects from aggregate displacement through firm reorganization to worker adjustment and distributional consequences. We ask how AI-driven automation alters the volume, composition, and geography of outsourced work, and how buyers and workers on both sides of the market respond. Understanding these dynamics matters because, unlike earlier waves of automation that primarily displaced workers in developed countries, generative AI directly targets tasks that had become globally tradable and that supported export-led growth in parts of the developing world. The adjustment costs of AI-driven automation may therefore be distributed across countries as well as across workers.

We address these questions using transaction data from a global digital labor platform specialized in cross-border outsourcing, where buyers, predominantly in developed countries, post tasks and workers in developing countries bid and complete contracts, providing a direct window into international trade in digital services. The setting offers three key advantages for studying how AI reshapes international outsourcing and its labor market consequences. First, we observe both sides of the market adjusting simultaneously: buyer reorganization and worker adaptation within the same institutional environment. Second, we track outsourcing behavior and worker adjustments from nearly a year before ChatGPT's release through two full years of diffusion. Third, the platform spans 2,700 narrowly defined skill categories with varying exposure to AI automation, enabling us to exploit rich cross-skill variation in a difference-in-differences design while tracing reallocation across the task space.

We combine these microdata with a novel, skill-specific measure of AI automation expo-

sure. Rather than relying on occupation-level proxies or realized adoption patterns, we directly elicit the capability of frontier generative AI models to perform each of the platform's 2,700 skills. The resulting exposure index captures within-occupation heterogeneity that coarser measures necessarily average away: two skills classified under the same occupation (e.g., "software developers") can differ substantially in automation potential depending on whether the underlying work is routine and codifiable or iterative and interaction-intensive. The index exhibits strong validity against both internal benchmarks and external measures. The capability-based measure identifies what AI can technologically perform, isolating the scope for automation from the organizational frictions that shape equilibrium adoption.

A task-based trade framework, building on models of task offshoring (Grossman and Rossi-Hansberg, 2008, 2012) and task-replacing technological change (Acemoglu and Restrepo, 2018, 2020), guides our empirical analysis. Buyers allocate tasks among domestic production, offshore outsourcing, and AI automation; workers with heterogeneous abilities choose which skills to supply and whether to invest in reskilling. By automating routine cognitive tasks, generative AI compresses the wage-productivity arbitrage that previously sustained offshore outsourcing. The framework generates four predictions: aggregate contraction of international outsourcing concentrates in developing-country destinations specialized in routine, codifiable tasks; buyer task consolidation into complex and AI-complementary projects; directed worker reskilling with differential success by ability; and rising within-skill concentration as threshold effects amplify inequality among those who remain.

We empirically investigate these hypotheses and document large effects propagating through both sides of the market. International outsourcing to developing countries declines sharply in high-exposure skills: at mean exposure, contract volume falls by approximately 35 percent, with the number of active buyers and workers declining in parallel. Displacement is concentrated in developing-country destinations; outsourcing to developed countries, where pre-AI comparative advantage involves complex and less codifiable tasks, shows no significant decline. Effects are not specific to a single AI product, with displacement deepening with each successive generation.

Beneath the aggregate contraction lies substantial reorganization on both sides of the market. Buyers fundamentally restructure international outsourcing by consolidating remaining work into fewer, larger, and more complex contracts. Total outsourcing value rises by 46 percent among continuing buyers while the number of contracts falls, implying a sharp increase in average contract value. This consolidation is accompanied by longer project durations, more detailed task specifications, and wider budget spreads, consistent with a shift toward work that is harder to automate and more difficult to price. Exposed buyers also increase their demand for projects explicitly referencing AI integration and deployment. These

patterns suggest that generative AI does not simply displace outsourcing but transforms its composition: buyers withdraw routine, codifiable tasks while retaining and expanding work that requires human judgment, coordination, and oversight.

On the supply side, workers face displacement but a subset actively adapts. Exposed workers complete 37 percent fewer contracts but earn 41 percent more per engagement, indicating a compositional shift toward higher-value work rather than uniform contraction. Beneath this average response lies substantial skill reallocation. High-exposure workers narrow the breadth of skills while simultaneously acquiring new competencies, with transitions directed toward AI-complementary, interaction-intensive tasks that retain value in an automated environment. Critically, this adaptation is highly uneven across the ability distribution: experienced and high-rated workers are significantly more likely to acquire new competencies, while less-established workers contract or exit. A gravity analysis of skill transitions confirms that reallocation follows technologically proximate pathways governed by switching costs and destination payoffs. The returns associated with adaptation are large: workers who successfully reskill into AI-complementary domains see significant earnings gains, while those who fail to reorient experience stagnant or declining earnings.

These heterogeneous responses generate sharp distributional consequences. The concentration of market opportunities rises by 22-24 percentage points in high-exposure skills, as top performers capture an increasing share of a shrinking market. This increase occurs primarily within skill categories rather than through reallocation across them, as high-exposure skills were already more concentrated than average before AI and their contraction slightly reduces aggregate concentration. Workers in exposed skills thus face a dual challenge: aggregate demand declines, and within the reduced market, opportunities concentrate among those best positioned to adapt. The above findings show that generative AI reshapes international outsourcing not through uniform contraction but through reorganization and selection: compressing the routine task trade that sustained developing-country comparative advantage while concentrating the remaining opportunities among workers best positioned to adapt.

Related literature. This paper contributes to the rapidly growing literature on generative AI and work by shifting the lens from domestic employment and within-job productivity to the reorganization of internationally traded digital services. Recent work quantifies the potential exposure of occupations to large language models and other AI technologies (Eloundou et al., 2024; Felten et al., 2019; Lou et al., 2023), measures AI usage through interaction or task-level data (Handa et al., 2025), and documents productivity effects through experimental and quasi-experimental designs (Noy and Zhang, 2023; Brynjolfsson et al., 2025; Dell'Acqua et al., 2023). We complement these contributions by moving beyond domestic

labor markets to study how generative AI reshapes the organization of production and trade across borders, documenting changes in realized outsourcing, task composition, and worker skill allocation in a global market for digital services.

A second strand uses online labor platforms, settings established as productive laboratories for studying contracting and labor market dynamics (Horton, 2010; Agrawal et al., 2016), to document how generative AI affects freelance work. Hui et al. (2024) find short-run displacement on a major platform, while Teutloff et al. (2025) show declining demand for AI-substitutable skills alongside growth in AI-complementary roles. We extend this work in three directions: we observe both sides of the market within a single setting rather than inferring one from the other; our platform specializes in cross-border outsourcing, enabling us to study how generative AI reshapes international trade in services and the labor markets of traditional outsourcing destinations; and we trace the full chain of adjustment, from aggregate displacement through buyer consolidation, directed reskilling, and rising concentration, linking demand-side restructuring to worker-level distributional consequences.

The paper also speaks to the broader literature on AI, automation, and labor markets. Earlier work emphasizes that AI affects a wide range of tasks without fully automating entire occupations (Brynjolfsson et al., 2018), while firm- and occupation-level studies find limited aggregate employment effects despite substantial within-firm reorganization (Acemoglu et al., 2022; Alekseeva et al., 2021). Our results show that muted domestic employment effects can coexist with large displacement in the international outsourcing markets through which routine cognitive work was previously traded, a margin largely absent from existing studies.

Finally, we contribute to the literature on trade in tasks and the organization of global production (Grossman and Rossi-Hansberg, 2008, 2012; Antràs and Chor, 2013; Antràs, 2015). While earlier work emphasizes how declining communication and coordination costs expand task offshoring, we document a force operating in the opposite direction: generative AI compresses the cross-border wage and productivity arbitrage that sustained routine task offshoring, withdrawing tasks from international trade not by reshoring them but by removing them from human labor. In contrast to prior studies of automation in developing countries, which focus on robotics and manufacturing (Faber, 2020; Artuc et al., 2023; DeStefano and Timmis, 2024), we study automation in digital services, a sector that emerged as a pathway to export-led growth in the global South, and trace how adjustment costs distribute across workers through directed reskilling and rising within-skill concentration.

The remainder of the paper proceeds as follows. Section 2 describes the cross-border online labor platform and the linked buyer-worker-skill transaction data. Section 3 describes the construction of our skill-level measure of exposure to generative AI. Section 4 presents a task-based conceptual framework formalized in Appendix A to motivate the empirical anal-

ysis and illustrate the mechanisms linking AI automation to outsourcing and worker adjustment. Section 5 documents aggregate effects of generative AI on cross-border outsourcing. Section 6 examines how exposed buyers reorganize outsourcing, focusing on changes in contract structure, project complexity, and task composition. Section 7 analyzes worker-side responses, including activity, reskilling, earnings dynamics, and the resulting distributional consequences. Section 8 concludes.

2 Empirical Setting and Data: Cross-Border Online Labor Market

Online labor marketplaces provide a unique micro-level view of trade in services by matching firms to workers worldwide for digitally deliverable tasks, a segment of trade that has expanded rapidly since the mid-2010s (World Bank, 2023).[1] Our empirical analysis uses data from a leading global online labor marketplace (OLM) that specializes in cross-border trade in digitally deliverable services.[2] Unlike many other platforms, which facilitate primarily domestic contracting or offer limited observability on realized transactions, this platform focuses on international outsourcing and fully mediates project-based contracts between firms and workers located in different countries. The data cover a broad range of activities organized into more than 2,700 skill categories, covering software development, web design, data analysis, marketing, writing and translation, business services, and more. Crucially, the platform records the universe of posted projects, bids, matches, and realized transactions, allowing us to observe detailed information on task content, contract terms, prices paid, work outcomes, and the geographic locations of both buyers and sellers.

These features also distinguish our data from job-posting datasets such as Burning Glass, which capture vacancies rather than realized employment relationships. The institutional structure of this platform allows us to observe both sides of the market and equilibrium outcomes, including selection into contracts, pricing, and the cross-country allocation of tasks. These data therefore provide an unusual lens on high-frequency adjustments in international outsourcing across extensive, intensive, and compositional margins that are largely unobservable in conventional trade and labor datasets.

[1] Two broad types of platforms operate in the gig economy: (i) location-based platforms (e.g., ride-sharing and local services) that require physical delivery, and (ii) remote work platforms (e.g., digitally deliverable professional services) where output can be produced and delivered online. Because generative AI primarily affects digitally produced content and code, we focus on the latter.

[2] We do not name the platform due to the terms of our data access agreement. Appendix B provides additional institutional details about the setting.

2.1 Key Features of the Platform

The defining characteristic of the platform in our study is its strong cross-border orientation. In our data, 91 percent of contracts involve a client and a worker located in different countries, with demand originating primarily in high-income economies, led by the United States, and supply concentrated in developing countries, particularly South Asia. This geographic structure reflects the platform's role as infrastructure for international trade in services rather than domestic gig work. It allows us to directly observe how AI-driven technological change affects the cross-country allocation of tasks, a margin that is central to debates on globalization and development but difficult to study using conventional data sources.

Compared to OLMs that primarily serve large and medium-sized firms with established procurement workflows, the platform attracts a broad spectrum of clients, from individual entrepreneurs and SMEs to government agencies and multinational corporations, reflecting its low posting costs and accessible bidding structure. This diversity is analytically valuable: it ensures that our data capture AI-driven reorganization across the firm size distribution, including among smaller buyers for whom outsourcing routine digital tasks may be the primary alternative to in-house production.

The platform organizes work into more than 2,700 skill categories. This fine-grained task structure enables measurement of AI exposure at the skill level and allows us to trace reallocation across narrowly defined activities, rather than across occupations or industries. Projects range from routine digital tasks, such as data entry or basic translation, to highly specialized services, including software architecture and machine learning implementation, generating substantial variation in task complexity and automation exposure.

Further, all contracts on the platform are mediated through a decentralized matching process that ensures comprehensive observability of outcomes. Clients post project listings specifying task requirements, budgets, and timelines, and workers submit proposals describing qualifications, experience, and compensation terms. Clients select among applicants based on observable characteristics such as price, skills, and prior performance. Once a contract is formed, the platform facilitates communication, milestone tracking, and payment, and records completion and performance feedback. This process generates a consistent and complete project-level record of outsourcing relationships and outcomes.

2.2 Data: Workers, Clients, and Transactions

We collect worker-, client-, and contract-level transaction data for all skill categories on the platform from January 2022 through December 2024. As in all online labor marketplaces, the population of registered users vastly exceeds the population of economically active par-

ticipants: a small core of workers generates nearly all completed contracts, earnings, and observable task allocation.[3] Our analysis focuses on this active core: the population for whom changes in AI capability translate into observable shifts in outsourcing volumes, contract terms, and skill composition.

To construct the analysis sample, we proceed in two steps. First, for each of the platform's 2,700 skill categories, we sample up to 500 workers who completed at least one contract in that skill before November 2022, following the platform's internal ranking system.[4] This captures nearly all workers with market visibility. The platform displays 10 workers per page; reaching the 500th worker requires navigating 50 pages of search results, well beyond the depth at which buyers typically engage, given the sharp decay of attention documented in digital marketplaces (Ursu, 2018). Given the granularity of the platform's skill taxonomy, over 95 percent of skill categories have fewer than 500 workers meeting our inclusion criterion; for these categories, our sample covers the entire active population. For larger categories, our sample captures the workers who appear in buyer searches and generate the transactions we observe. Because individual workers typically supply multiple skills, the same worker may appear in multiple skill-level samples. After de-duplicating across skills, the procedure yields approximately 93,000 unique individuals. In the pre-period (January–October 2022), these workers accounted for approximately 95 percent of all completed contracts on the platform, consistent with the extreme concentration characteristic of digital labor markets (Horton, 2010; Agrawal et al., 2016).

Second, for each worker in the cohort, we collect the complete history of completed projects visible on their profile between January 2022 and December 2024, yielding approximately 300,000 contracts. The cohort is fixed at entry: workers are included based on pre-November 2022 activity and followed regardless of whether they remain active, exit the platform, or change their skill composition in the post-period.[5] This design allows us to capture both intensive-margin adjustment among continuing workers and extensive-margin displacement through exit, within a single cohort observed over the full sample period.

The final dataset consists of linked client-worker-skill transaction microdata for cross-border services trade. Worker profiles provide location, experience, education, and certifica-

[3]Throughout the paper, we interpret outcomes as changes in observed outsourcing activity within the platform. Our analysis captures reorganization of digitally traded services as reflected in realized contracts; it is not designed to measure firms' in-house labor adjustments or total labor demand outside the platform.

[4]The ranking incorporates completed contracts and client ratings and determines the order in which workers appear in buyer search results.

[5]Fixing the cohort ensures that changes in outsourcing volumes, earnings, or task composition are not mechanically driven by confounding shifts in platform composition. To the extent that AI deters entry by workers who would have joined absent the technology, our estimates of aggregate displacement may be conservative.

tions (when available), prior job history, and client ratings. Client records include country, hiring history, industry classifications, and ratings. Project-level data contain titles and descriptions, associated skill categories, contract status, budget ranges, hours worked, observed payments, and multidimensional performance ratings covering quality, communication, timeliness, and professionalism.[6] The structure of the data allows us to trace high-frequency adjustments in outsourcing behavior, task allocation, and worker outcomes in response to generative AI. Because we observe realized contracts rather than intended demand, we can detect both immediate and persistent changes in client and labor behavior that are largely invisible in aggregate labor market and trade statistics.

3 Measuring Skill-Specific Exposure to Generative AI Automation

We construct a skill-specific measure of AI exposure that captures heterogeneity in generative AI capabilities across the platform's skill categories. The measure is designed to reflect how firms evaluate automation potential when making outsourcing decisions: by assessing whether frontier models can perform the work associated with a given skill.

Definition and measurement. We define AI exposure as the extent to which generative AI can perform the work associated with a given skill with minimal human input, thereby reducing the need for human labor. This capability-based concept measures the technological frontier for automation as perceived by market participants, rather than realized adoption, abstracting from firm-specific organizational frictions that shape equilibrium deployment. The distinction matters: our objective is to identify skills where AI *could* substitute for human contractors, not skills where substitution has already occurred.

For each skill, we elicit assessments of generative AI capability using structured prompts. Each prompt provides the model with three pieces of information drawn from the platform's taxonomy: the skill name, its broad professional category (e.g., "Websites, IT & Software"), and the platform's definition of the skill. The model is then asked to rate automation feasibility:

> "On a scale from 0 to 1, how well can generative AI perform a task related to the skill [SKILL NAME] compared to a human, with minimal human input?"

where 0 indicates that AI cannot meaningfully perform the skill and 1 indicates full human-

[6] Appendix B provides additional details on the data collection and variable definitions.

level capability.[7] The baseline index uses GPT-3.5, the model underlying the release of ChatGPT in November 2022. This choice is deliberate: it aligns the exposure measure with the capabilities available when firms first began adjusting outsourcing decisions. To the extent that subsequent models (GPT-4, GPT-4o) shift the automation frontier, our baseline measure provides a conservative estimate of exposure. We verify in Appendix C that the exposure rankings are highly stable across model generations (Spearman $\rho = 0.63$ between GPT-3.5 and GPT-4).

Validation. Table 1 reports skills with the highest and lowest AI exposure. The ranking exhibits strong face validity. High-exposure skills center on text-based, rule-governed tasks such as programming, translation, proofreading, data entry for which LLMs can perform core activities with minimal human input. Low-exposure skills involve obsolete technologies and highly specialized domains that require tacit knowledge or context-specific judgment.

The ranking also reveals substantial within-field heterogeneity. For example, while general programming ranks among the most exposed skills, legacy technologies (e.g., COBOL) appear near the bottom, reflecting limited training data and narrower model applicability. This pattern suggests that exposure is driven not by technical sophistication per se, but by the extent to which tasks can be specified through codifiable instructions and evaluated using standardized criteria.

We assess robustness to both alternative constructions and external benchmarks. First, we construct an internal alternative that replaces full automation with AI-assisted task execution, asking how well generative AI can "assist an individual in performing the tasks" rather than perform them autonomously. The two measures are highly correlated ($\rho = 0.87$), indicating that the automation and augmentation frontiers largely overlap at the current state of the technology; the distinction is one of degree rather than kind. We use the automation-focused measure as our baseline because it more directly captures the margin relevant to outsourcing decisions: whether firms can replace offshore labor with AI, not whether AI can make that labor more productive.

Second, we benchmark against three prominent occupation-level indices: the α (assistance) and β (automation) measures from Eloundou et al. (2024), who classify O*NET tasks by susceptibility to large language models, and the Anthropic AI Exposure Index, which captures realized AI usage patterns across occupations. To implement these occupation-level measures in our skill-based setting, we map each skill category to corresponding O*NET tasks and occupations through a multi-stage matching procedure.[8] Correlations between

[7] Appendix C documents the complete system role instructions and prompt template.

[8] Appendix C.3 details the matching procedure, reports the full correlation structure, and discusses sources of divergence between skill-level and occupation-level measures.

our measure and the external indices range from 0.52 to 0.67, indicating meaningful alignment while leaving variation that reflects within-occupation skill heterogeneity our measure is designed to capture.[9] Appendix D.5 re-estimates our baseline specifications using each alternative measure; results are qualitatively similar, with the magnitudes attenuated when using coarser occupation-level indices.

Relation to existing measures. Our approach differs from prior methodologies along three dimensions. First, it operates at the level of skills and captures the margin along which reallocation occurs in digital outsourcing markets where contracts specify narrowly defined tasks and technological change is more likely to reshape the composition of work within fields than to uniformly expand or contract entire occupations. Second, it measures technological capability rather than realized usage, separating what AI can do from how firms choose to deploy it.[10] Third, it reflects how economic agents probe AI in practice through direct model interaction, making the measure a behavioral proxy for firms' beliefs about automation potential.

A potential concern is that a language model evaluates its own capabilities. However, this mirrors firm behavior: managers assess automatability by testing models' ability to perform specific skills. To the extent that self-assessment errors are not correlated with true exposure, measurement error attenuates estimated effects toward zero. Moreover, the 2,700 skills span domains over which no group of human evaluators possesses uniform expertise. Human-based assessment would inevitably embed discipline-specific priors that are not commensurate across fields. LLM elicitation thus provides a practical alternative to infeasible comprehensive human assessment.

Complementary skill measures. Beyond exposure to automation, we construct three additional measures to capture reallocation across skill types. First, we construct an alternative *assistance* exposure measure based on AI-assisted task execution rather than full automation. The prompt asks how well generative AI can assist an individual in performing the tasks associated with a given skill, capturing settings in which human input remains necessary. Appendix D.5 reports the analysis using this alternative measure.

Second, we measure *AI intensity* based on the share of contracts within each skill that involve AI-related work: for example, tasks explicitly referencing AI integration, prompt engineering, model deployment, or workflow automation. This intensity reflects the skill's scope for integrating generative AI, rather than vulnerability for substitution.

[9] Figure A.2 illustrates the heterogeneity of skill-level exposure scores within select O*NET occupations.

[10] Measures based on application logs or user prompts are informative about equilibrium deployment but may conflate capability with adoption decisions. Our approach isolates the automation frontier.

Third, we measure *interactive intensity*: the share of projects within each skill whose descriptions reference communication, client interaction, creativity, or coordination. This captures dimensions of work relying on interpersonal judgment that remain less amenable to automation even when the underlying task is digitally deliverable.

These measures allow us to distinguish three margins of adjustment: displacement from automatable tasks, expansion into AI-complementary work, and reallocation toward interaction-intensive services. In the empirical analysis, we use these measures to assess whether observed changes in outsourcing and task composition reflect substitution away from automatable work, upgrading toward AI-complementary technical skills, and reallocation toward interactive tasks that remain human-intensive.

4 Conceptual Framework

This section develops a task-based framework linking generative AI to changes in international service outsourcing. The framework builds on canonical models of task-based production and offshoring (Grossman and Rossi-Hansberg, 2008, 2012) and on theories of task-replacing technological change (Acemoglu and Restrepo, 2018, 2020). Appendix A provides a formal model; here we present the key mechanisms, derive testable predictions, and discuss conditions under which alternative channels could modify or reverse the baseline results.

We characterize tasks along two dimensions: exposure to AI automation, namely, the extent to which generative AI can perform the task autonomously, and complexity, namely, the degree to which the task requires coordination, customization, or judgment. We assume these dimensions are negatively correlated in the population of digitally tradable tasks.[11]

These task characteristics interact with cross-country differences in wages and productivity to determine pre-AI outsourcing patterns. Developing countries' lower wages confer a comparative advantage in routine, codifiable tasks where labor cost is the primary determinant of sourcing decisions. More complex tasks favor workers in developed countries, whose productivity advantages in coordination, communication, and judgment offset higher wages. Prior to generative AI, tasks with high automation exposure are therefore disproportionately outsourced to developing countries, while developed-country contractors specialize in lower-exposure, higher-complexity work (Grossman and Rossi-Hansberg, 2008, 2012).

[11] This assumption reflects the empirical regularity that highly automatable tasks such as routine coding, translation, and standardized content generation tend to be modular and codifiable, while tasks that resist automation, such as system architecture, strategic consulting, and creative direction, tend to require tacit knowledge and adaptive judgment. The negative relationship between codifiability and complexity is a standard feature of task-based models (Autor et al., 2003; Acemoglu and Restrepo, 2018) and of the relational contracting literature (Antràs, 2015).

The public release of large language models in late 2022 constituted a discrete improvement in the quality and availability of AI tools capable of performing text-based, code-based, and design-adjacent tasks. For buyers, this improvement reduces the effective cost of performing high-exposure tasks through automation. Tasks that were previously outsourced because developing-country wages undercut in-house alternatives now become candidates for algorithmic production. Because high-exposure tasks are concentrated in developing-country outsourcing, this substitution disproportionately displaces contractors in developing countries. Outsourcing to developed countries, concentrated in complex tasks that remain below the automation threshold, is comparatively insulated.

This substitution mechanism is the central channel in our framework, but it is not the only one through which generative AI may affect outsourcing. AI may also augment worker productivity, enabling faster or higher-quality output rather than replacing human labor entirely. If augmentation dominates substitution, outsourcing could increase as AI-assisted workers become more competitive. A third possibility is task expansion: AI may create entirely new categories of work—prompt engineering, model deployment, AI-human workflow design—that generate fresh demand for outsourced labor. Appendix A formalizes all three channels and derives conditions under which each dominates.[12] The empirical analysis tests whether the net effect is displacement or expansion, rather than imposing the substitution channel ex ante.

Under the substitution channel, as automatable tasks exit the outsourced set, the composition of remaining work shifts toward tasks that are more complex, interdependent, and difficult to specify ex ante. Such tasks benefit from relational contracting: they require coordination across subtasks, adaptation over time, and accumulated client-specific knowledge. Combined with fixed contracting costs, this compositional shift leads buyers to consolidate outsourcing into fewer, larger contracts with contractors capable of handling broader project scope and sustained interaction. The model predicts that contract values, project durations, and task specifications all increase among continuing buyers.

On the supply side, generative AI alters the relative value of skills. Skills tied to automatable tasks decline in value, while skills that facilitate coordination, integration, and client interaction, including AI-complementary and interactive skills, appreciate as buyers reorganize toward complex projects. Workers reallocate in response, with transitions shaped by the expected return to destination skills and adoption frictions that increase with the distance between origin and destination skills.[13] Critically, adaptation capacity varies across workers.

[12]The augmentation and task-expansion channels are not merely theoretical alternatives; they are empirically relevant margins we examine in Sections 6 and 7. Finding net displacement in our data would indicate that substitution dominates augmentation and task creation over the period we study.

[13]Appendix A shows that this structure generates a gravity-like pattern in which worker flows are localized

High-ability workers face lower reskilling costs due to greater experience and learning capacity (Acemoglu, 1998). This heterogeneity implies that the same technological shock produces divergent outcomes: experienced and high-rated workers expand into AI-complementary domains, while less established workers are more likely to contract their portfolios or exit.

These forces jointly increase the concentration of market opportunities through three reinforcing channels. First, demand contraction induces exit among workers in high-exposure tasks, particularly those with lower ability. Second, as buyers consolidate into complex, multi-skill projects, the capability requirements for winning contracts rise, rationing out less versatile workers. Third, high-ability workers reskill more effectively into AI-complementary roles, capturing disproportionate shares of growing categories. The framework makes a sharper prediction about the source of rising concentration. The three channels above—selection, threshold effects, and differential adaptation—all operate within skills, unambiguously increasing inequality among workers competing in the same domain. The cross-skill channel is ambiguous in sign: it depends on whether high-exposure skills were more or less concentrated than average before AI.

The framework yields four testable hypotheses that structure our empirical analysis. Appendix A formalizes the mechanisms and derives each prediction.

Prediction 1: Task displacement. Outsourcing volume declines more sharply in skills with higher AI automation exposure, with the contraction concentrated among contractors in developing countries whose pre-AI comparative advantage lay in routine, codifiable tasks (Section 5).

Prediction 2: Contract consolidation. Continuing buyers shift toward fewer, larger, and more complex contracts, with increases in average contract value, project duration, and AI-complementary task content (Section 6).

Prediction 3: Directed reskilling. Workers in high-exposure skills reallocate toward AI-complementary and interaction-intensive activities. Reskilling is concentrated among high-ability workers (Section 7.2).

Prediction 4: Rising concentration. Within-skill concentration of earnings and contracts increases in high-exposure skills, driven by selection, rising capability thresholds, and differential adaptation among continuing workers (Section 7.4).

We examine these predictions in the following sections, exploiting variation in AI exposure across skills, buyers, and workers, and transaction-level data that allow us to observe realized

along technologically proximate pathways but directed toward higher-value activities.

contracts, task reallocation, and skill adjustment at high frequency.

5 Aggregate Effects on International Outsourcing

We begin by examining the first prediction of the framework: whether generative AI leads to differential contraction in international outsourcing across skills with varying automation exposure, and whether this contraction is concentrated in developing-country destinations. Aggregate skill-level analysis provides a natural starting point before we trace how the effects are mediated through buyer reorganization (Section 6) and worker adjustment (Section 7).

Empirical specification. Our empirical strategy is a difference-in-differences design that exploits two sources of variation: (i) the timing of the arrival and diffusion of generative AI, and (ii) cross-skill heterogeneity in the technological scope for automation, measured by our time-invariant exposure index $Exposure_i$ constructed in Section 3. Intuitively, we compare changes in outsourcing before and after the onset of the generative-AI era for skills that are more versus less exposed to automation.

We estimate the following baseline specification at the skill-week level:

$$y_{it} = \gamma_i + \lambda_t + \beta\big(Exposure_i \times PostGPT_t\big) + \varepsilon_{it}, \tag{1}$$

where i indexes skill subcategories and t indexes weeks. The outcome y_{it} denotes measures of outsourcing activity in skill i at time t, including the total value and number of completed contracts. The indicator $PostGPT_t$ equals one for weeks after the onset of widespread access to generative AI (our baseline cutoff is the public release of ChatGPT), and zero beforehand. Skill fixed effects γ_i absorb time-invariant differences across skills, including baseline demand, typical contract size, and market thickness. Week fixed effects λ_t flexibly control for common shocks to the platform such as seasonality, aggregate demand fluctuations, and macroeconomic conditions. Standard errors are clustered at the skill level to account for serial correlation within skills.

The coefficient β captures whether outsourcing changes differentially after the introduction of generative AI in skills with greater automation exposure. Because $Exposure_i$ is time-invariant, this is a continuous-treatment difference-in-differences design: identification comes from differential changes over time across skills as a function of their ex ante exposure to AI automation. Appendix D.1 reports estimates using exposure quartile bins rather than the continuous measure, allowing for non-linear dose-response patterns.

The key identifying assumption is that, absent generative AI, outsourcing in high- and low-exposure skills would have evolved along parallel trajectories conditional on skill and

time fixed effects. We assess this assumption along two dimensions. First, we test for differential pre-trends directly using the event-study specification below:

$$y_{it} = \gamma_i + \lambda_t + \sum_{\tau \neq -1} \beta_\tau \big(Exposure_i \times \mathbb{1}\{t = \tau\} \big) + \varepsilon_{it}, \tag{2}$$

where τ indexes months relative to November 2022. Pre-period coefficients $\{\beta_\tau\}_{\tau<0}$ provide a direct test for differential pre-trends; post-period coefficients trace the dynamic evolution of exposure-dependent effects.

Second, a remaining concern is that shocks contemporaneous with ChatGPT's release, such as macroeconomic tightening or platform-specific changes, could differentially affect skills in ways correlated with automation exposure. For such shocks to confound our estimates, they would need to affect skills in proportion to their AI automation scores, a specific correlation structure that generic demand or supply shocks are unlikely to produce.[14]

We estimate the baseline specification on a panel of skill-week observations from January 2022 to December 2024. Our main analysis focuses on outsourcing to developing countries, which accounts for 93 percent of cross-border transactions on the platform; we then estimate effects separately for developed-country destinations to test for the asymmetry predicted by the framework.

Aggregate effects. Table 2 reports the effects of generative AI on international outsourcing at the skill–week level. Consistent with the first prediction of the framework, we find large and statistically significant declines in outsourcing concentrated in skills with greater exposure to AI automation. Column (1) reports estimates for log total contract value: the coefficient on $Exposure_i \times PostGPT_t$ is -0.820, implying that a skill with a mean exposure (0.69) experienced a decline in outsourcing value of approximately 43 percent.[15] Column (2) shows a corresponding decline in contract volume of 35 percent at mean exposure. These estimates indicate a substantial contraction in both the scale and frequency of outsourced work in skills where AI can most readily substitute human labor. Columns (3) and (4) decompose this contraction into participation margins.[16] Both the number of buyers posting contracts

[14] Our design does not require that firms or workers immediately adopt generative AI. We interpret the release of ChatGPT as a capability shock that expands the technological frontier. The difference-in-differences design identifies how outsourcing responds to changes in what AI can perform, rather than to realized adoption. Even if take-up is gradual, reorganization of production ultimately reflects changes in technological capability.

[15] Note that the estimates capture relative effects: the contraction in high-exposure skills is measured relative to low-exposure skills. If workers reallocate from high- to low-exposure skills and activity in low-exposure skills subsequently rises, the measured gap overstates the differential effect but understates aggregate displacement. We examine these reallocation dynamics directly in Section 7.

[16] Sections 6 and 7 explore the intensive margins among continuing buyers and workers.

and the number of workers completing contracts decline significantly with exposure. At mean exposure, market participation falls by roughly 34–35 percent on both sides.

Prediction 1 further implies that displacement is concentrated in developing-country destinations, where pre-AI comparative advantage lay in the routine, codifiable tasks most susceptible to automation. Figure 4 evaluates this prediction by comparing the effects on developing- and developed-country destinations. The contraction in outsourcing to developing countries is large and statistically significant across all outcomes as discussed above. In contrast, the response in developed countries is close to zero and statistically insignificant. This asymmetry is consistent with the framework: developing-country contractors specialized in high-exposure tasks bear the brunt of displacement, while developed-country outsourcing, concentrated in complex tasks below the automation threshold, is comparatively insulated.[17]

Figures 1 and 2 examine the dynamics underlying these aggregate effects. Across all outcomes—total contract value, number of projects, active buyers, and active workers—event-study coefficients are flat and statistically indistinguishable from zero prior to the diffusion of generative AI, providing no evidence of differential pre-trends between high- and low-exposure skills. The exposure gradient emerges gradually following the release of GPT-3.5 and becomes markedly more negative after the release of subsequent models. The common timing of these shifts across value, volume, and participation margins indicates that the contraction in outsourcing coincides with discrete improvements in AI capability rather than with pre-existing differences across skills.

Appendix D reports various specification checks. First, the results are stable to alternative functional forms, including the inverse hyperbolic sine transformation to accommodate zeros and the use of exposure bins instead of a continuous measure. The estimates are also robust to more demanding fixed effects, including category-by-time fixed effects that absorb differential trends across broader skill groups. Third, we take into account major model releases across multiple AI providers, including OpenAI's GPT-4 and GPT-4o, Anthropic's Claude 3, and Google's Gemini, and find significant incremental displacement effects. Finally, the estimates are qualitatively similar but attenuated using our alternative measure based on AI-assisted production rather than full automation and existing occupation-based indices, including the Eloundou et al. (2024) α and β measures and the Anthropic AI Exposure Index, consistent with the differences in what each measure is designed to capture. While the alternative indices provide useful validation of the aggregate results, quantifying adjustments by market participants requires exposure defined at the level of platform-specific

[17]The imprecision of the developed-country estimate may also partly reflect the small volume of outsourcing to these destinations. We therefore interpret this result as consistent with destination asymmetry. Appendix D.7 provides finer evidence, showing that the contraction is largest in Bangladesh and India, the countries that specialized most heavily in routine, codifiable digital services prior to generative AI.

skills that map directly to observed contracts, task bundles, and worker portfolios.[18]

6 Buyer and Task Reorganization

The aggregate results establish that outsourcing contracts differentially in high-exposure skills. But aggregate skill-level regressions can mask substantial reorganization in what buyers continue to contract and how they structure outsourced work. The second prediction of the framework implies that as routine tasks exit the outsourced set, the residual work shifts toward complex, interdependent activities that favor consolidation into fewer, larger contracts with broader scope and greater coordination requirements. This section tests that prediction directly by moving from skill-week to buyer-level analysis, tracking how individual buyers adjust contract structure, project scope, and task composition as a function of their pre-period exposure to generative AI.

For each outcome, we estimate

$$y_{bt} = \gamma_b + \lambda_{ct} + \beta\big(\text{Exposure}_b \times \text{PostGPT}_t\big) + \varepsilon_{bt}, \tag{3}$$

where γ_b are buyer fixed effects capturing buyers' time-invariant characteristics and λ_{ct} are source-country-week fixed effects absorbing source country-specific labor demand and supply shocks. Buyer Exposure_b is constructed using the buyer's pre-GPT outsourcing portfolio and represents the maximum automation exposure of all skills outsourced by buyer b prior to November 30, 2022. This measure captures the buyer's ex ante vulnerability to AI-driven substitution. We use the maximum rather than the mean because the framework predicts that buyers reorganize once their highest-exposure tasks cross the automation threshold, regardless of the average exposure of their portfolio.[19] We examine buyer reorganization along two margins: (i) contract structure and spending, and (ii) task composition and complexity.

6.1 Outsourcing Structure and Budgets

Table 3 presents estimates of Equation (3) for buyers who remain active after the introduction of ChatGPT. The sample restriction to continuing buyers is deliberate: having documented

[18] We also show in Appendix D.7 that the magnitude of the response varies systematically across destination countries. The contraction is largest in Bangladesh and India, countries that, prior to generative AI, specialized most heavily in routine, codifiable digital services. In contrast, effects are smaller and less precisely estimated for countries such as Egypt, where baseline outsourcing volumes are lower and skill composition is more heterogeneous. This pattern suggests that rather than uniformly affecting all destinations, generative AI disproportionately erodes demand for tasks that formed the core of comparative advantage in certain developing-country labor markets.

[19] Results are qualitatively similar using value-weighted mean exposure across pre-period contracts.

significant buyer exit in the aggregate analysis (approximately 34 percent at mean exposure, Table 2, Column 3), we now ask how *surviving* buyers restructure their outsourcing.

The results reveal systematic reorganization among high-exposure buyers. Column (1) shows that total outsourcing spending increases by approximately 46 percent at mean buyer exposure. Column (2) indicates that this increase does not reflect more frequent contracting: the number of contracts declines by approximately 12 percent. Column (3) confirms the implication: average contract value rises sharply, by approximately 65 percent. Exposed buyers consolidate into fewer, larger contracts rather than scaling back proportionally across all margins. This consolidation is consistent with the framework's prediction that, as routine tasks exit the outsourced set, remaining work shifts toward complex, interdependent activities that benefit from bundling into broader-scope projects.[20]

The increase in total spending among continuing buyers contrasts with the aggregate decline documented in Section 5. This divergence reflects composition: the aggregate effect combines intensive-margin adjustment (continuing buyers spend more per contract) with extensive-margin exits (high-exposure buyers are more likely to leave the platform entirely). Among those who remain, consolidation into fewer, larger contracts produces higher total spending despite reduced contract frequency. This pattern of simultaneous exit at the extensive margin and upgrading at the intensive margin is a distinctive prediction of the framework that would not arise under uniform contraction.

To assess whether high- and low-exposure buyers were on differential trajectories prior to generative AI, we estimate an event-study analog of Equation (3). Figure 3 reports the results. Event-study coefficients for contract counts and average contract value are flat in the pre-period and diverge sharply after November 2022, providing no evidence that the observed reorganization reflects pre-existing trends.[21]

The shift toward larger contracts is accompanied by changes in buyers' stated willingness to pay. Table 4 examines budget parameters that buyers specify when posting contracts. Column (1) shows that the maximum posted budget increases by approximately 32 percent

[20] An alternative interpretation is that firms substitute away from this platform toward other intermediaries or off-platform contracting. While we cannot directly observe off-platform activity, three observations argue against pure disintermediation. First, the contraction in outsourcing is sharply increasing in task-level automation exposure, which is difficult to reconcile with platform-specific substitution. Second, we document systematic changes in the structure of remaining projects in Section 6.2, including consolidation, higher budgets, and broader skill requirements, which align with the substitution of automatable tasks rather than simple platform exit. Third, worker-side evidence in Section 7 shows directed reskilling toward AI-complementary skills rather than diffuse reallocation, consistent with technology-driven reorganization.

[21] Appendix D.6 explores heterogeneity in buyer responses and finds evidence of selective reorganization rather than uniform contraction: exposure shifts outsourcing toward higher-value work among buyers with the technical scope to adapt, while buyers with greater pre-period scale are more likely to substitute away from outsourced production.

at mean exposure; Column (2) documents a 22 percent rise in the minimum budget. The increase in the ceiling exceeds that of the floor: Column (3) shows that the budget spread, i.e., the difference between maximum and minimum posted budgets, widens significantly.

The widening spread is informative about the nature of reorganization. For routine, codifiable tasks, buyers can specify narrow budget ranges because market prices are well established. The fact that budget uncertainty increases differentially in high-exposure categories suggests that remaining projects are less standardized, involving tasks for which appropriate compensation is harder to benchmark ex ante. This pattern is consistent with a shift toward complex, multi-dimensional work whose scope and difficulty are more uncertain at the time of posting.

6.2 Task Complexity and AI-Complementary Projects

The consolidation patterns documented above are consistent with the framework's prediction but could in principle reflect repricing of equivalent work (e.g., if reduced worker supply in high-exposure skills increases equilibrium wages) rather than a genuine shift in the nature of outsourced tasks. To distinguish between these interpretations, we examine direct measures of project complexity observable in the transaction data.

We focus on four proxies for project scope and composition, each aggregated to the buyer-week level. The first is *turnover time*: the number of days between contract posting and completion review, capturing the realized duration of the buyer-worker engagement.[22] The second is *description length*: the word count of the project posting. More complex projects, particularly those requiring coordination across multiple skills or deliverables, typically demand more detailed specifications of requirements, scope, and expected outputs. The third is *AI-related share*: the fraction of a buyer's contracts in a given week that explicitly reference generative AI tools or tasks.[23] This measure captures whether exposed buyers reorient demand toward work that complements generative AI, such as deployment, customization, evaluation, or human oversight of automated systems, rather than work that AI replaces. The fourth is *interactive share*: the fraction of contracts requiring communication, client interaction, creativity, or interpersonal coordination, capturing whether reorganization shifts outsourcing toward relationship-intensive tasks.[24]

[22]Longer turnover time could, in principle, reflect matching frictions rather than project complexity. We interpret increases as evidence of complexity for two reasons: they coincide with longer project descriptions (suggesting greater scope) and with rising contract values (suggesting buyers are paying for more work, not waiting longer for the same work).

[23]We classify a project as AI-related if the description contains terms related to large language models, specific model names (e.g., GPT, Claude, LLaMA), or AI-adjacent tasks such as prompt engineering, model fine-tuning, and AI integration. Appendix C.4 provides a list of keywords.

[24]Appendix C.4 details the classification procedure for both measures.

Table 5 reports estimates of Equation (3) for these outcomes. Column (1) shows that turnover time increases by approximately 52 percent at mean buyer exposure, indicating that projects take substantially longer to complete. Column (2) documents a corresponding increase in description length: the tasks that exposed buyers continue to outsource require more detailed specification of requirements and deliverables. The joint increase in duration and specification detail points to greater underlying project complexity rather than repricing of equivalent tasks or matching delays.

Column (3) examines the AI-related project share. Exposed buyers are significantly more likely to post projects referencing generative AI tools following November 2022. While the coefficient is modest in absolute terms, it is economically meaningful relative to the near-zero baseline prevalence of such references in the pre-period. This result suggests that buyers previously reliant on automatable tasks partially redirect demand toward work that complements AI—projects requiring human expertise to deploy, fine-tune, or supervise automated systems rather than tasks that AI performs autonomously.

Column (4) examines the interactive project share. The point estimate is positive but not statistically significant. This null result may reflect that interaction-intensive tasks were already concentrated among developed-country contractors less affected by the AI shock, or that interpersonal coordination is bundled within complex projects rather than contracted separately. The above results indicate that buyer reorganization operates primarily through increased technical complexity and AI integration rather than through a general shift toward interpersonal work.

An important interpretive distinction is whether the observed shifts reflect active reorganization whereby buyers restructure their outsourcing toward new types of work, or compositional changes whereby simple projects disappear while complex projects continue unchanged. The evidence favors active reorganization: Tables 3 and 4 show that continuing buyers increase total spending and average contract value, not merely that small contracts are subtracted from a stable portfolio. The increase in AI-related content, which had negligible prevalence before November 2022, further indicates that buyers are adding new types of projects rather than passively retaining residual activity.

6.3 AI Intensity and Outsourcing Responses

The preceding analysis characterizes responses among buyers whose outsourced tasks are susceptible to AI substitution. We now examine a distinct margin of heterogeneity: buyers' pre-existing *AI intensity*, the share of a buyer's pre-November 2022 contracts that involved

AI-related work.[25] This measure captures organizational readiness to integrate generative AI into production, as opposed to vulnerability to displacement. The distinction is economically meaningful: a firm that was already outsourcing machine learning tasks before November 2022 faces a different optimization problem than a firm that was outsourcing data entry.

Table 6 reports estimates of Equation (3) with AI intensity replacing automation exposure as the treatment variable. The pattern is strikingly different. Column (1) shows that total outsourcing value declines for AI-intensive buyers following November 2022. However, Column (2) reveals a significant increase in the number of contracts, while Column (3) shows a sharp decline in the mean contract value. AI-intensive buyers respond to generative AI not by consolidating but by decomposing outsourcing into a larger number of smaller, more modular contracts.

The opposing patterns across the two buyer-level measures—consolidation among high-automation-exposure buyers, decomposition among AI-intensive buyers—are economically informative and help distinguish the substitution and augmentation channels. Automation exposure governs whether tasks exit the outsourced set: as AI substitutes for routine work, residual outsourcing consolidates into complex bundles. AI intensity governs how production is reorganized once firms actively integrate AI as an input: these buyers outsource complementary human tasks around automated processes, such as specialized evaluation, integration, troubleshooting, or task-specific customization, leading to finer task decomposition and smaller contract units.

This divergence has implications for interpreting the aggregate results. The overall contraction in outsourcing reflects the dominance of the substitution channel across the platform. But among the subset of organizationally prepared firms, generative AI generates new demand for modular, complementary human inputs that interface with automated systems. The coexistence of both patterns within the same marketplace highlights that generative AI reshapes the organization of outsourcing, not merely its volume, and that the direction of reorganization depends on whether firms are positioned to use AI as a substitute for external labor or as a complement to it.

[25] AI intensity is measured using the same keyword classification applied to pre-period contracts. A buyer with high AI intensity had already been outsourcing AI-related tasks before ChatGPT; a buyer with high automation exposure had been outsourcing tasks that AI can now perform autonomously. The two measures are conceptually distinct and only weakly correlated in the data, as shown in Appendix C.4.

7 Worker Adjustment, Skill Reallocation, and Distribution of Opportunities

While the earlier sections document how firms reorganize outsourcing in response to advances in generative AI, an equally important margin of adjustment operates on the supply side of the market. A central advantage of our setting is that we observe the complete history of realized contracts and the skills attached to each contract at high frequency, enabling direct measurement of within-worker adjustments, a margin that is typically unobservable in studies relying on occupation- or industry-level employment data. Guided by the framework in Section 4, we organize the analysis around four margins: (i) how exposed workers adjust their level of market activity, (ii) how they restructure their skill portfolios and adopt new competencies, (iii) how earnings respond to different modes of adjustment, and (iv) how these forces jointly shape the distribution of opportunities across workers.

7.1 Worker Adjustments

We examine how workers in more automatable tasks adjust their activity and earnings following the diffusion of generative AI. For worker i in week t, we estimate

$$y_{it} = \gamma_i + \lambda_{ct} + \beta \left(\text{Exposure}_i \times PostGPT_t\right) + \varepsilon_{it}, \tag{4}$$

where y_{it} denotes outcomes including the number of jobs completed, total earnings, and mean job value. Exposure_i is a time-invariant measure of worker exposure, defined as the maximum automation score across all skills utilized by worker i in tasks performed prior to November 30, 2022. The logic parallels the buyer-level measure: the framework predicts that workers begin adjusting once their most automatable skill crosses the substitution threshold, regardless of the average exposure of their portfolio. Worker fixed effects γ_i absorb time-invariant differences in ability, experience, and specialization, while country-time fixed effects λ_{ct} control for country-specific demand conditions and platform-wide shocks.[26] Standard errors are clustered at the worker level.

Table 7 examines intensive-margin adjustment among workers who remain active on the platform after November 2022. Having documented significant extensive-margin exit in the aggregate analysis (approximately 35 percent fewer active workers at mean exposure, Table 2, Column 4), we now ask how continuing workers adjust their activity. Column (2) reports

[26] Country-by-week fixed effects ensure that comparisons are made among workers in the same country at the same time, so that differential trends in outsourcing demand across developing countries do not confound the exposure gradient.

estimates for the log number of completed contracts per worker-week: at mean exposure (0.52), contract volume declines by approximately 37 percent relative to unexposed workers. Demand for workers specializing in automatable skills contracts sharply.

Column (1) examines total earnings. The point estimate is negative but not statistically significant. The imprecision of the earnings estimate, combined with the significant decline in contract volume, implies that changes in contract composition partially offset reduced transaction frequency. Column (3) confirms this: at mean exposure, continuing workers earn approximately 41 percent more per contract in the post-period. The contrast between declining contract volume and rising contract value mirrors the buyer-side consolidation documented in Section 6.1: as buyers bundle residual work into fewer, more complex projects, workers who continue to win contracts are compensated at higher rates per engagement.

This pattern is inconsistent with uniform demand contraction, which would reduce both volume and value. It is also inconsistent with a pure supply-side story in which worker exit mechanically raises prices for remaining workers: if the effect were driven solely by reduced supply, we would expect earnings to rise proportionally with exit rates, whereas the compositional shift toward complex, multi-skill contracts (documented below) indicates genuine reorganization of the work being performed.

We additionally examine whether generative AI affects the quality of work performed by re-estimating Equation (4) using client ratings across multiple dimensions (communication, professionalism, experience, and overall quality). Average ratings exhibit little change with exposure. However, allowing for heterogeneity by pre-period performance reveals that quality improvements are concentrated among workers with lower baseline ratings (Figure 5), consistent with AI tools disproportionately augmenting less experienced workers—a pattern documented in other settings (Brynjolfsson et al., 2025).[27]

7.2 Skill Reallocation and Directed Reskilling

The preceding analysis shows that continuing workers perform fewer but higher-value contracts following the diffusion of generative AI. A central question is whether this adjustment reflects passive selection whereby workers simply retain their most remunerative tasks while losing lower-value work, or active reallocation of the skills. We address this distinction by examining how the skill composition of workers' contracts changes in response to AI exposure. Our measures capture market-validated competencies— skills that buyers actually pay

[27]This result should be interpreted cautiously: it could also reflect compositional change if lower-rated workers who remain active are positively selected relative to lower-rated workers who exit.

for, reflecting the equilibrium reallocation of work across skill categories.[28]

Skill portfolio restructuring. We characterize each worker's skill portfolio along four dimensions: (i) the total number of distinct skills, (ii) the number of newly adopted skills not appearing in any of the worker's pre-period contracts, (iii) the number of pre-existing skills retained from the pre-period, and (iv) the average number of skills per contract. These measures allow us to distinguish portfolio contraction (fewer skills, no replacement) from active restructuring (displaced skills replaced with new competencies).

Table 8 reports estimates of Equation (4) for these outcomes. Column (1) shows that the total number of distinct skills in a worker's contracts declines by 1.98 at mean exposure. In isolation, this could reflect demand contraction in high-exposure skills without any worker-initiated adjustment. However, Columns (2) and (3) reveal pronounced compositional change beneath this aggregate decline: high-exposure workers significantly increase the number of newly adopted skills while simultaneously reducing the number of pre-period skills appearing in their contracts. The opposing signs indicate active replacement, i.e., workers entering new skill categories as they exit exposed ones, rather than passive narrowing.

Column (4) provides complementary evidence from the contract level. High-exposure workers increasingly complete projects requiring a broader set of skills within a single engagement. This shift toward multi-skill contracts is the supply-side counterpart of buyer consolidation documented in Section 6: as buyers bundle previously separable tasks into complex projects, workers who remain competitive must supply complementary skills jointly rather than in isolation.[29]

Columns (5) and (6) characterize the skills toward which workers reallocate. Column (5) shows that high-exposure workers significantly increase their share of projects referencing generative AI tools such as deployment, fine-tuning, prompt engineering, or oversight of automated systems.[30] Column (6) documents an even larger increase in the share of projects involving interactive skills: communication, coordination, creativity, or client-facing judgment. Reallocation is directed toward activities that complement generative AI: either by interfacing directly with AI systems or by supplying interpersonal and contextual inputs that resist automation. The combination of AI-related and interaction-intensive adoption

[28] Profile-listed skills may include aspirational or outdated competencies; contract-based skills reflect what the market values at a given point in time. For our purposes, realized allocations are the object of interest.

[29] A potential concern is that observed portfolio changes reflect relabeling with workers tagging existing work with different skill labels, rather than genuine reallocation. This concern is mitigated because the platform maintains a centrally administered skill taxonomy that workers cannot redefine. Moreover, Columns (5) and (6) validate reallocation using the textual content of completed projects: newly adopted skills coincide with measurable changes in project descriptions, which is difficult to reconcile with purely cosmetic relabeling.

[30] AI-related and interaction-intensive projects are classified using the keyword procedures described in Section 6.2 and detailed in Appendix C.4.

suggests that workers are not simply fleeing exposed categories but repositioning toward the emerging complementarities created by the same technology that displaces their prior work.

Figure 6 illustrates these patterns by visualizing observed transition paths for workers initially specialized in Adobe Photoshop, a high-exposure skill. Adaptation is highly structured: a substantial share of workers continue completing Photoshop contracts or closely related work, indicating that not all activity in high-exposure domains is displaced. Among those who reallocate, transitions occur primarily toward technologically proximate skills: some toward technical skills such as HTML/CSS and JavaScript, where generative AI assists with code generation but human input remains essential for integration, and others toward adjacent creative design or marketing tasks requiring judgment, customization, and client interaction.

Appendix D.8 characterizes these patterns systematically using a gravity-style specification motivated by the model in Appendix A. Skill-pair relatedness strongly predicts transition flows, confirming that workers move primarily along lower-friction margins. Conditional on relatedness, workers reallocate toward skills with greater pre-period payoffs and higher interaction intensity, consistent with the model's prediction that both switching costs and expected returns govern the direction of reallocation.

Who reskills? Next, we examine whether the above skill reallocation is broadly shared across exposed workers or concentrated among workers with greater capacity to adapt. We focus on two margins: the acquisition of new skills in realized contracts and the disappearance of previously used skills from a worker's contract portfolio. Specifically, we define indicators for whether a worker's contracts in a given week include at least one skill not present in any pre-period contract (new skill addition), include an AI-related or interaction-intensive skill for the first time, or no longer include a skill that appeared in pre-period contracts (skill shedding). We then allow these responses to vary with worker characteristics measured prior to November 2022.

Table 9 reports estimates. Panel A interacts the post-GPT exposure effect with pre-period experience, measured as the log number of contracts completed before November 2022. The baseline coefficient on Post-GPT $\times$ Exposure is negative: on average, higher exposure is associated with fewer skill additions and greater skill shedding. However, the interaction with experience reverses this pattern. Conditional on exposure, more experienced workers are significantly more likely to add new skills—including AI-related skills—and significantly less likely to lose pre-period skills from their contract portfolios. Panel B documents analogous heterogeneity by baseline performance, measured using pre-period client ratings. Higher-rated workers are significantly more likely to add interaction-intensive skills and substantially

less likely to shed previously used skills.

These results indicate that active skill reallocation is concentrated among workers with the strongest pre-existing market positions. This heterogeneity has a natural interpretation through the framework presented in Section 4: experienced and high-rated workers face lower effective switching costs because their broader skill bases provide more proximate pathways to AI-complementary domains, and their established reputations reduce the credentialing frictions that constrain entry into new skill markets. The concentration of adaptation capacity among already-advantaged workers provides a micro-foundation for the distributional consequences examined in Section 7.4.

7.3 Earnings and Skill Adjustments

We next examine how different modes of skill adjustment correlate with worker earnings in the post-period. These results are associational: workers who successfully reorient their skill portfolios likely differ from those who do not on unobservable dimensions including motivation, learning ability, and access to information. We therefore interpret the estimates as characterizing the earnings correlates of different adjustment strategies rather than as causal returns to reskilling.[31]

Table 10 reports estimates of Equation (4) augmented with interactions between the post-period exposure effect and indicators for different types of skill adjustment. The baseline interaction between exposure and the post-GPT indicator is small and insignificant across all columns, suggesting that exposure alone does not mechanically translate into earnings changes. Earnings trajectories instead diverge with the mode of adjustment.

Column (1) shows that adding new skills is associated with approximately 21 percent higher earnings at mean exposure. The association is considerably stronger when reallocation targets specific domains: adding AI-related skills is associated with approximately 63 percent higher earnings (Column 2), and adding interaction-intensive skills also yields significant gains (Column 3). Column (4) examines workers who shed previously used skills: earnings decline significantly when shedding occurs without compensating additions, but are fully offset among workers who simultaneously add AI-related or interaction-intensive skills. Column (5) confirms this pattern using net switching intensity, i.e., the balance of skill additions over subtractions, which is positively associated with earnings.

Two features of these results merit emphasis. First, the large premium associated with AI-related skill additions, in contrast to the limited aggregate transition flows into AI des-

[31] Identifying causal returns would require exogenous variation in skill adjustment conditional on exposure (for instance, differential access to training programs or randomized information about emerging skill demands).

tinations documented in Appendix D.8, is consistent with the gravity model's prediction that high switching costs can sustain earnings differentials in equilibrium. The workers who manage to overcome entry barriers into AI-complementary domains capture substantial rents precisely because few displaced workers can follow. Second, the finding that skill shedding without compensating reorientation predicts earnings losses, while shedding accompanied by new additions does not, suggests that the earnings consequences of displacement depend critically on whether workers actively redirect effort rather than passively contract their market participation.

7.4 Concentration of Market Opportunities

The preceding sections document how generative AI reshapes the volume and composition of outsourced work and how workers adjust through skill reallocation. We now examine how these individual-level responses aggregate into changes in the distribution of market opportunities. Generative AI can affect concentration through two distinct channels: a *within-skill* channel, where concentration rises within existing skill categories as some workers exit or lose market share to top performers, and a *cross-skill* channel, where activity reallocates toward skills that differ in their baseline concentration levels.

The framework in Section 4 generates an unambiguous prediction for the within-skill channel: concentration rises through threshold effects (buyer consolidation raises capability requirements for winning contracts), selection (differential exit across the ability distribution), and differential adaptation (high-ability workers reskill more effectively, as documented in Table 9). The cross-skill channel is ambiguous *a priori*: if high-exposure skills were more concentrated than average before November 2022, their contraction reduces aggregate concentration through composition; the reverse would increase it. We now test whether concentration rises in practice and through which channels.

Let s_{it} denote skill i's share of platform earnings at time t, and C_{it} denote within-skill concentration, measured as the share of earnings or contracts captured by the top five workers in skill i at t. Aggregate concentration is $\bar{C}_t = \sum_i s_{it} C_{it}$. Changes in aggregate concentration decompose as:

$$\Delta \bar{C} = \underbrace{\sum_i \bar{s}_i \Delta C_i}_{\text{within-skill}} + \underbrace{\sum_i \bar{C}_i \Delta s_i}_{\text{cross-skill}} + \text{interaction}, \tag{5}$$

where $\bar{s}_i$ and $\bar{C}_i$ denote pre-November 2022 values. The within-skill component captures concentration changes within categories weighted by their baseline importance; the cross-skill component captures reallocation of activity across categories weighted by their baseline concentration levels.

We further decompose the within-skill component to distinguish the role of worker turnover from rising inequality among continuing workers. Let C_{it}^{pre} denote the top-5 share in skill i at time t computed only among workers who were active in that skill prior to November 2022. Then:

$$\Delta C_i = \underbrace{(C_{it} - C_{it}^{\text{pre}})}_{\text{participation}} + \underbrace{(C_{it}^{\text{pre}} - C_{i,\text{baseline}})}_{\text{within-active}}. \tag{6}$$

The participation component captures how entry and exit alter concentration; the within-active component captures whether top performers among continuing workers gain market share over time. As derived in Appendix A, the sign of the participation component provides a diagnostic for the nature of selection: a positive coefficient indicates exit concentrated among low performers (mechanically raising the top-5 share), while a negative coefficient suggests that some top performers also exit from some pre-period skill domains, potentially because they possess superior outside options.

We estimate the effect of AI exposure on each component using Equation (1), regressing each concentration measure on skill-level exposure interacted with the post-period indicator.

Table 11, Panel A reports estimates for within-skill concentration. The coefficients are positive and significant: at mean exposure, the top-5 earnings share increases by approximately 27 percentage points and the top-5 contract share by 26 percentage points.

Panel B examines the cross-skill component. The coefficients are negative but economically small, contributing approximately one percentage point to the change in aggregate concentration. High-exposure skills were slightly more concentrated than average before November 2022, so their contraction after generative AI's arrival (documented in Section 5) marginally reduces aggregate concentration through composition. However, this channel is an order of magnitude smaller than the within-skill effect: rising platform-wide concentration is driven overwhelmingly by what happens *within* skill categories, not by the reallocation of activity across them.

Panel C decomposes the within-skill effect. The participation component is negative: worker turnover in high-exposure skills actually reduces measured concentration, implying that exiting is not confined to the bottom of the ability distribution. This result is consistent with the evidence in Section 7.2 and the gravity estimates in Appendix D.8. Top performers in high-exposure skills are disproportionately likely to reskill successfully toward AI-complementary domains, effectively exiting their original skill category. The negative participation component thus reflects two forces: (i) high performers transitioning out of exposed skills into AI-complementary work, and (ii) new entrants or workers expanding from adjacent categories diluting concentration. Both reduce measured concentration relative to what would prevail if only the pre-period workforce were counted.

The within-active component is positive and significant: among workers who were active in a skill before November 2022 and remain active in that same skill, top performers capture an increasing share of opportunities over time. This is the channel through which the framework's predicted mechanisms—rising capability thresholds and differential adaptation—manifest empirically. As buyers consolidate into complex, multi-skill projects (Section 6.1), the workers best equipped to deliver these work gain disproportionately within their skill categories.

Panel D translates skill-level coefficients into implied effects on platform-wide concentration: the net effect is approximately 24 percentage points for earnings and 22 for contracts, driven almost entirely by the within-skill channel. The concentration of adaptation capacity among high-ability workers, that is, those who are both more likely to reskill out of exposed categories and more likely to gain market share within the categories where they remain, suggests that generative AI amplifies pre-existing differences in worker positioning rather than creating new ones.

8 Conclusion

This paper investigates how generative AI reshapes international outsourcing of digital services, using transaction-level data from a global online labor marketplace to jointly analyze buyer reorganization and worker adjustment within a single market. The combination of high-frequency contract data, granular skill classifications, and observation of both sides of each transaction allows us to provide a unified account of how an automation technology shock alters the organization and geography of digital service production.

The empirical analysis yields four main findings. First, generative AI produces large, asymmetric displacement: contract volume in developing-country outsourcing declines by approximately 35 percent at mean exposure, especially in countries specialized most heavily in routine, codifiable digital services. Second, buyers do not simply reduce spending: they reorganize and consolidate demand into fewer, higher-value, more complex engagements as AI substitutes for routine tasks while raising the returns to bundled, interaction-intensive work. Third, workers in high-exposure skills actively reallocate their skill portfolios, but the capacity to do so is sharply uneven: experienced and high-rated workers expand into AI-complementary domains, while less-established workers contract or exit. A gravity-style analysis confirms that reallocation follows technologically proximate pathways directed toward higher-payoff destinations. Fourth, these heterogeneous responses generate rising within-skill concentration as top performers capture a growing share of a shrinking pool of opportunities.

Our analysis opens several questions that the present setting cannot fully resolve. We

study a single platform; whether the effects documented here generalize to traditional offshoring arrangements, where contracts are governed by longer-term relationships and switching costs are higher, is an important empirical question, though the mechanisms we identify operate through forces that are not platform-specific. We observe the first two years following ChatGPT's release, a period in which AI capabilities expanded rapidly across successive model generations. Whether the displacement documented here represents a transitional phase as workers, buyers, and institutions develop new forms of human-AI complementarity or the early stage of a deeper structural shift will depend on the trajectory of AI capabilities relative to the pace of human adaptation. Our evidence is informative but not dispositive: the estimates suggest that switching costs are a binding constraint on worker adjustment, while the cumulative model-release analysis shows displacement deepening rather than stabilizing with each new generation of AI.

The findings carry three broader implications for how economies absorb rapid technological change. First, the speed of adjustment is striking, raising questions about whether existing adjustment mechanisms are calibrated to the relevant timescale. Second, the distributional consequences are self-reinforcing: the workers best positioned to adapt are those already advantaged, while those most exposed face the highest barriers to reskilling. The gravity estimates formalize this tension: high switching costs sustain earnings differentials even when destination returns are large, suggesting that targeted interventions to reduce entry barriers into AI-complementary domains may be necessary to broaden access to adjustment. Third, the international dimension is central to the welfare calculus. The countries driving AI innovation are not those bearing the largest adjustment costs; what appears as productivity gains from the perspective of developed-country buyers manifests as displacement in the developing-country labor markets that supply these services.

References

Acemoglu, Daron, "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics*, 1998, *113* (4), 1055–1089.

_ **and David Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, pp. 1043–1171.

_ **and Pascual Restrepo**, "Artificial intelligence, automation, and work," in "The economics of artificial intelligence: An agenda," University of Chicago Press, 2018, pp. 197–236.

_ **and** _ , "Artificial Intelligence and Jobs," *Journal of Economic Perspectives*, 2020, *34* (3), 30–50.

_ **, David Autor, Jonathon Hazell, and Pascual Restrepo**, "Artificial intelligence and jobs: Evidence from online vacancies," *Journal of Labor Economics*, 2022, *40* (S1), S293–S340.

Agrawal, Ajay, John J. Horton, Nicola Lacetera, and Elizabeth Lyons, "Digitization and the Contract Labor Market: A Research Agenda," *Management Science*, 2016, *62* (7), 1953–1972.

Alekseeva, Liudmila, José Azar, Mireia Gine, Sampsa Samila, and Bledi Taska, "The demand for AI skills in the labor market," *Labour Economics*, 2021, *71*, 102002.

Antràs, Pol, *Global Production: Firms, Contracts, and Trade Structure*, Princeton University Press, 2015.

_ **and Davin Chor**, "Organizing the Global Value Chain," *Econometrica*, 2013, *81* (6), 2127–2204.

Artuc, Erhan, Paulo Bastos, and Bob Rijkers, "Robots, tasks, and trade," *Journal of International Economics*, 2023, *145*, 103828.

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

Baldwin, Richard, *The Globotics Upheaval: Globalization, Robotics, and the Future of Work*, Oxford University Press, 2019.

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond, "Generative AI at Work," *The Quarterly Journal of Economics*, 2025, *140* (2), 889–942.

_ , Tom Mitchell, and Daniel Rock, "What can machines learn, and what does it mean for occupations and the economy?," *AEA papers and proceedings*, 2018, *108*, 43–47.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R Lakhani, "Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, 2023, (24-013).

DeStefano, Timothy and Jonathan Timmis, "Robots and export quality," *Journal of Development Economics*, 2024, *168*, 103248.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs are GPTs: Labor Market Impact Potential of LLMs," *Science*, 2024, *384* (6702), 1306–1308.

Faber, Marius, "Robots and reshoring: Evidence from Mexican labor markets," *Journal of International Economics*, 2020, *127*, 103384.

Felten, Edward W, Manav Raj, and Robert Seamans, "The occupational impact of artificial intelligence: Labor, skills, and polarization," *NYU Stern School of Business*, 2019.

Grossman, Gene M and Esteban Rossi-Hansberg, "Trading tasks: A simple theory of offshoring," *American Economic Review*, 2008, *98* (5), 1978–1997.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Task Trade between Similar Countries," *Econometrica*, 2012, *80* (2), 593–629.

Handa, Kunal, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, Sarah Heck, Jared Mueller, Jerry Hong, Stuart Ritchie, Tim Belonax, Kevin K. Troy, Dario Amodei, Jared Kaplan, Jack Clark, and Deep Ganguli, "Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations," 2025.

Horton, John J., "Online Labor Markets," in Amin Saberi, ed., *Internet and Network Economics*, Vol. 6484 Springer 2010, pp. 515–522.

Hui, Xuan, Oren Reshef, and Luofeng Zhou, "The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market," *Organization Science*, 2024.

Lou, Bowen, Hongshen Sun, and Tianshu Sun, "Gpts and labor markets in the developing economy: Evidence from china," *Available at SSRN 4426461*, 2023.

Noy, Shakked and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science*, 2023, *381* (6654), 187–192.

Teutloff, Ole, Johanna Einsiedler, Otto Kässi, Fabian Braesemann, Pamela Mishkin, and R. Maria del Rio-Chanona, "Winners and Losers of Generative AI: Early Evidence of Shifts in Freelancer Demand," *Journal of Economic Behavior & Organization*, 2025, *235*, 106845.

Ursu, Raluca M., "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions," *Marketing Science*, 2018, *37* (4), 530–552.

World Bank, *Digital Trade for Development* 2023.

Figures and Tables

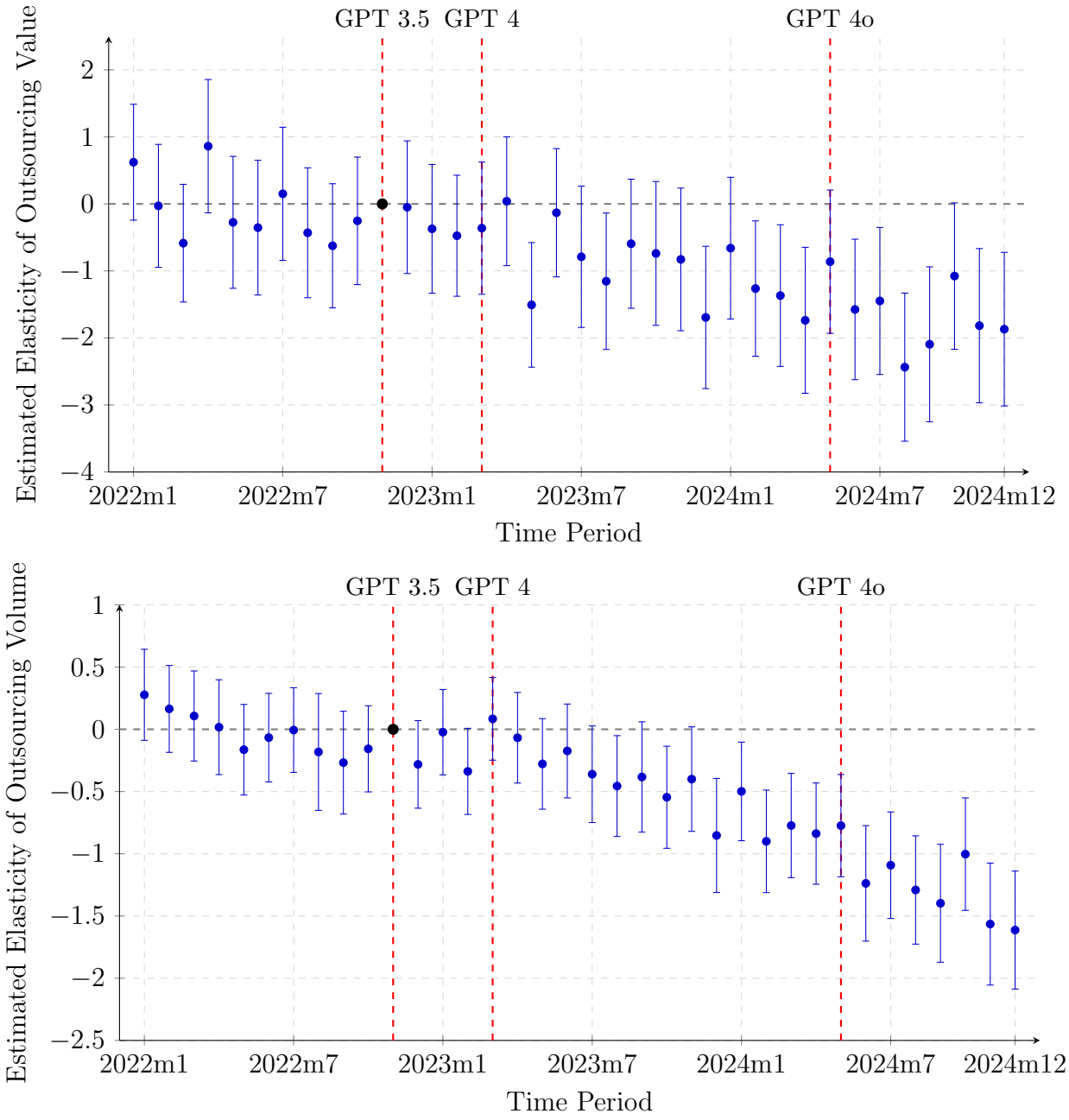


Figure 1: Event Study: Dynamic Effects of Generative AI on Outsourcing

Notes. These figures present event study estimates of the dynamic effects of generative AI model releases on international outsourcing. Each point represents the coefficient estimate for a given month relative to the baseline period, with 95% confidence intervals shown. Vertical dashed red lines indicate the release dates of GPT-3.5 (November 30, 2022), GPT-4 (March 14, 2023), and GPT-4o (May 13, 2024). The horizontal dashed gray line indicates zero effect. Standard errors are clustered at the skill level. The top panel shows the elasticity of total outsourcing value, while the bottom panel shows the elasticity of the number of contracts.

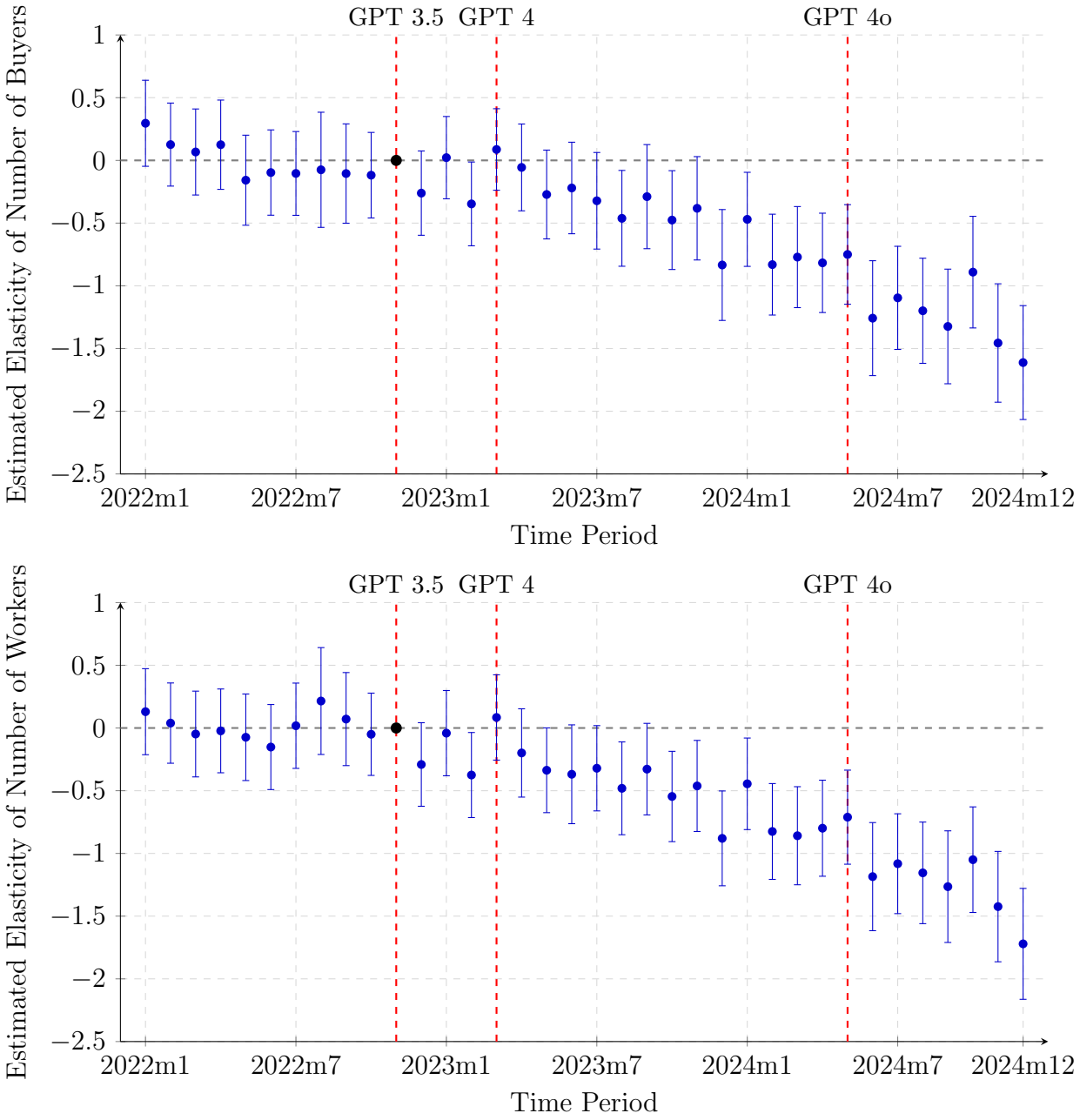


Figure 2: Event Study: Dynamic Effects of Generative AI on Buyers and Workers

Notes. These figures present event study estimates of the dynamic effects of generative AI model releases on international outsourcing. Each point represents the coefficient estimate for a given month relative to the baseline period, with 95% confidence intervals shown. Vertical dashed red lines indicate the release dates of GPT-3.5 (November 30, 2022), GPT-4 (March 14, 2023), and GPT-4o (May 13, 2024). The horizontal dashed gray line indicates zero effect. Standard errors are clustered at the skill level. The top panel shows the elasticity of the number of buyers, while the bottom panel shows the elasticity of the number of workers.

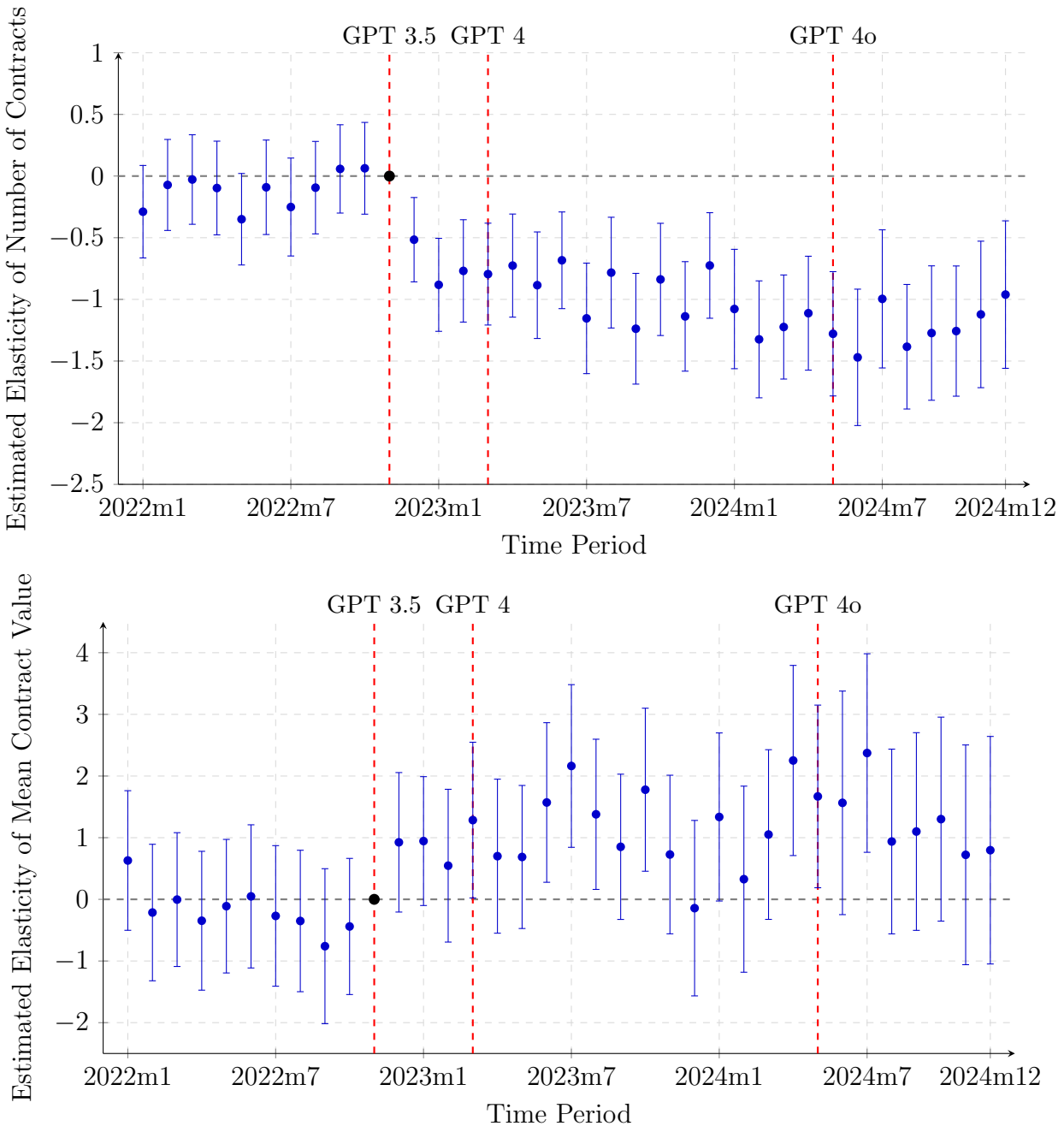


Figure 3: Event Study: Dynamic Effects of Generative AI on Buyer-Level Contract Volume and Average Contract Value

Notes. These figures present event study estimates of the dynamic effects of generative AI model releases on international outsourcing. Each point represents the coefficient estimate for a given month relative to the baseline period, with 95% confidence intervals shown. Vertical dashed red lines indicate the release dates of GPT-3.5 (November 30, 2022), GPT-4 (March 14, 2023), and GPT-4o (May 13, 2024). The horizontal dashed gray line indicates zero effect. Standard errors are clustered at the skill level. The top panel shows the elasticity of the buyers' number of contracts, while the bottom panel shows the elasticity of the buyers' average contract value.

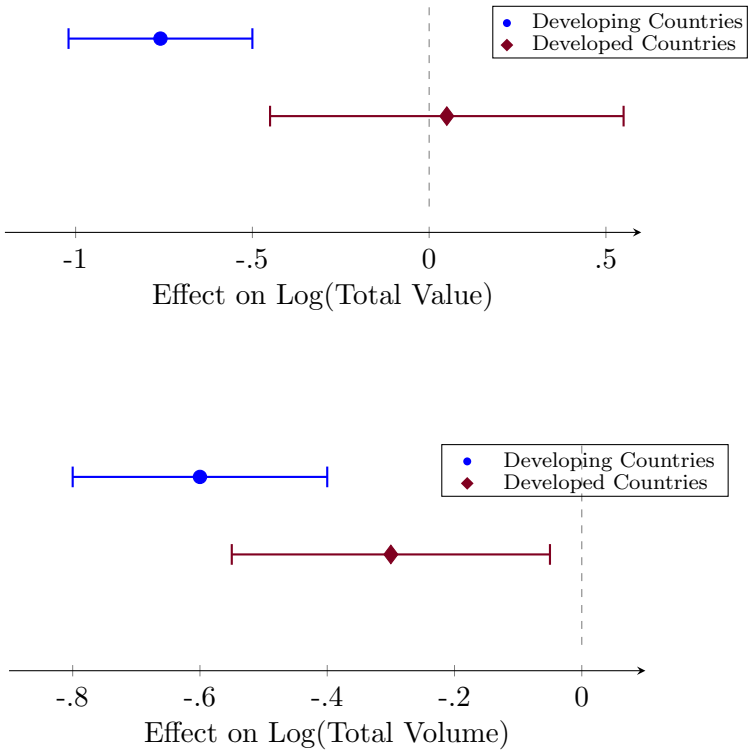


Figure 4: Aggregate Effects: Developing vs Developed Countries

Notes: The figure compares the estimated coefficients from Equation (1) for developed and developing economies, along with their 95% confidence intervals. The top panel shows the impact of generative AI on the total value of contracts outsourced to each group of countries, while the bottom panel illustrates the effect on the volume of contracts outsourced.

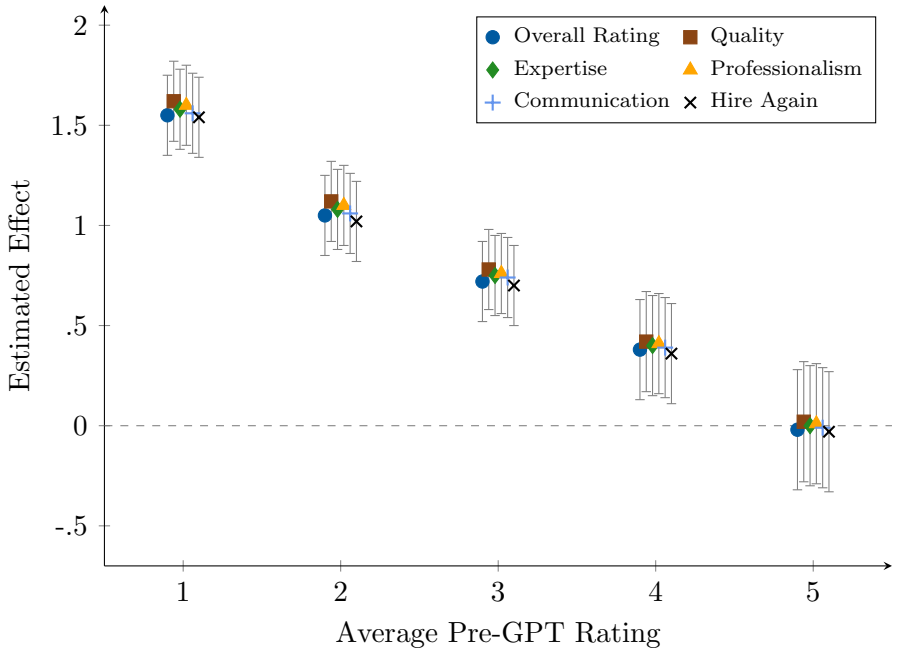


Figure 5: The Heterogeneous Effects of Generative AI across Worker Quality

Notes: The figure plots the estimated effects of generative AI on worker ratings based on the levels of their average ratings before the introduction of ChatGPT, along with their 95% confidence intervals. Ratings include overall rating and ratings on quality, expertise, professionalism, communication, and the likelihood to hire again.

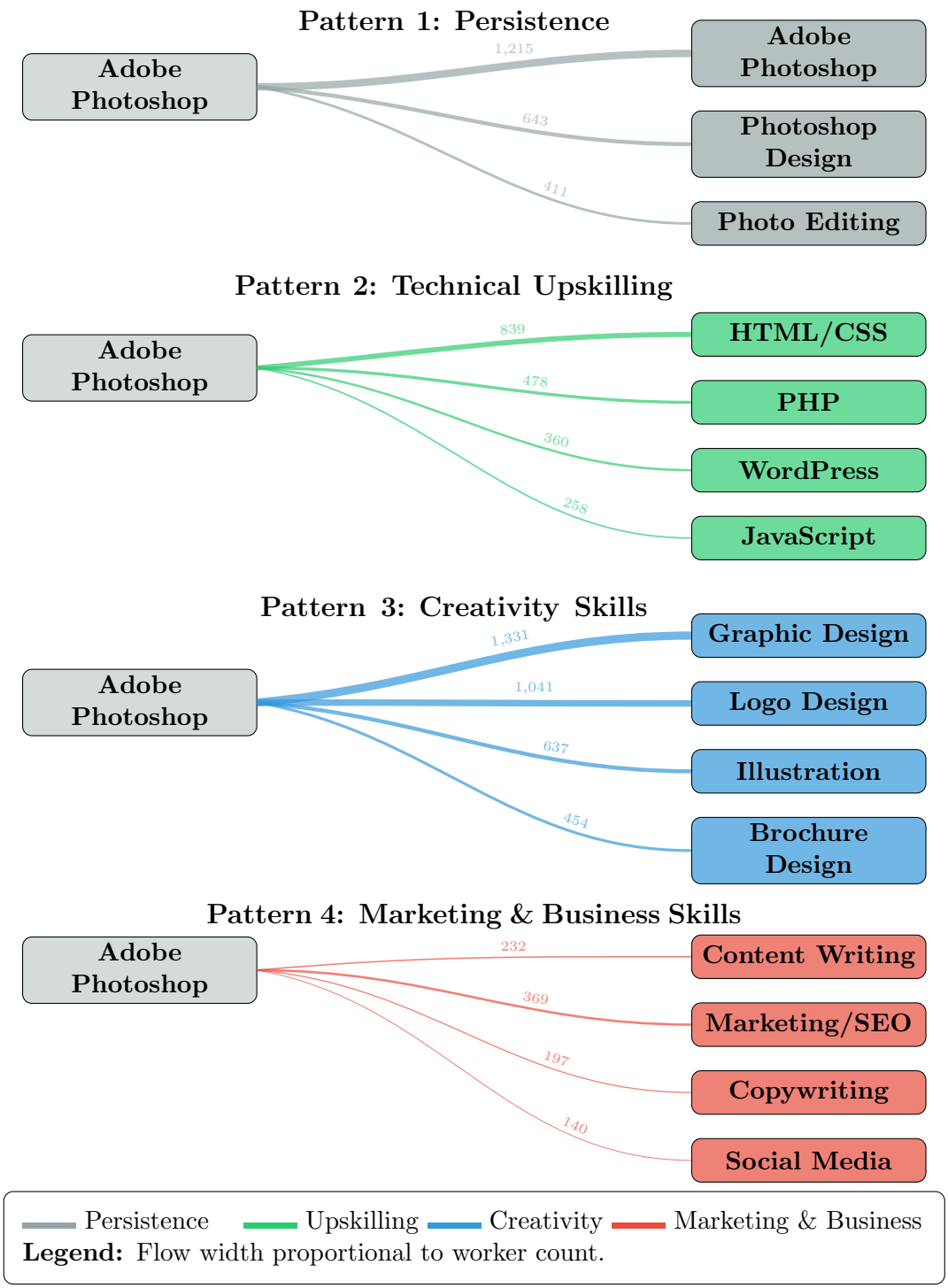


Figure 6: Worker Adaptation Strategies from High-AI-Exposure Skills

Notes. This figure maps skill transitions from Adobe Photoshop (a skill with high AI automation exposure) following ChatGPT's November 30, 2022 release, organized by adaptation strategy. *Pattern 1 (Persistence):* 2,269 workers remain in Photoshop-related skills despite high AI substitution risk. *Pattern 2 (Technical Upskilling):* 1,935 workers transition to complementary technical skills (HTML, PHP, JavaScript) that involve system architecture and logic—areas where AI serves as a tool rather than a substitute. *Pattern 3 (Creativity Skills):* 3,463 workers shift to related design skills (Graphic Design, Logo Design). *Pattern 4 (Marketing & Business Skills):* 938 workers move to marketing and content strategy tasks requiring client interaction, business judgment, and creative strategy. These patterns reveal that while some workers persist within high-exposure domains, a substantial portion (Patterns 2-4) reposition toward technical, creative, or client-facing tasks.

Table 1: Most and Least Generative AI Automatable Skills

Most Automatable Skills			Least Automatable Skills		
Rank	**Skill**	**Score**	**Rank**	**Skill**	**Score**
1	Google Search	0.948	2718	Anthropology	0.314
2	Software Engineering	0.943	2719	Attic Access Ladders Making	0.314
3	English Spelling	0.930	2720	Christmas Planning	0.312
4	Programming	0.929	2721	Genealogy	0.308
5	English Grammar	0.928	2722	Abnormal Psychology	0.306
6	Coding	0.925	2723	Forth (Programming Language)	0.306
7	Word Processing	0.924	2724	Feng Shui	0.302
8	English (UK) Translation	0.921	2725	Alto Flute	0.295
9	HTML	0.916	2726	Palm	0.295
10	Adobe Photoshop	0.914	2727	Soldering	0.287
11	Android Development	0.913	2728	COBOL	0.282
12	Proofreading	0.905	2729	ASHSP	0.279
13	CMS	0.905	2730	MMORPG	0.274
14	Mathematics	0.905	2731	Building Certification	0.258
15	Writing	0.905	2732	Symbian	0.254
16	English (US) Translation	0.904	2733	Asbestos Removal	0.250
17	Typing	0.904	2734	Papiamento	0.244
18	Google Chrome	0.903	2735	Sculpturing	0.233
19	ChatGPT Usage	0.902	2736	Google Buzz	0.225
20	Web Development	0.902	2737	Fire Safety Planning	0.224
21	Copy Typing	0.902	2738	ECPay	0.200
22	Firefox	0.901	2739	Fashion Modeling	0.192
23	Photoshop	0.899	2740	Google Wave	0.191
24	Web Applications	0.896	2741	MySpace	0.144
25	Drawing / Artist	0.895	2742	Flash Mob	0.132

Notes. This table reports the 25 skills with the highest and lowest exposure to generative AI automation among the platform-defined skill categories. Exposure scores are constructed by eliciting assessments from GPT-3.5 of the extent to which generative AI can perform the tasks associated with each skill with minimal human input. Scores are averaged over 100 independent, stateless prompt iterations using fixed decoding parameters and normalized to lie between 0 and 1. Higher values indicate greater automation exposure. Scores should be interpreted in relative rather than absolute terms. Skill labels correspond to the platform's fixed taxonomy and do not necessarily map one-to-one to standard occupational classifications.

Table 2: Aggregate Effects on International Outsourcing

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Buyers)	(4) Log(Workers)
Post-GPT × Exposure	-0.820***	-0.608***	-0.601***	-0.612***
	(0.147)	(0.089)	(0.085)	(0.079)
Constant	6.652***	1.491***	1.457***	1.409***
	(0.071)	(0.043)	(0.041)	(0.038)
Skill FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
R^2	0.637	0.887	0.892	0.884
Observations	88,800	90,045	90,045	90,045

Notes. The dependent variables are measured at the skill-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a skill-level exposure measure capturing GPT's assessed ability to substitute for workers in that skill, normalized to the unit interval. All specifications include skill fixed effects and week fixed effects. Standard errors, clustered at the skill level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Continuing Buyers' Outsourcing Margins

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Mean Value)
Post-GPT × Exposure	1.095***	-0.360***	1.454***
	(0.217)	(0.044)	(0.212)
Constant	4.232***	0.216***	4.010***
	(0.083)	(0.017)	(0.081)
Buyer FE	Yes	Yes	Yes
Source Country × Week FE	Yes	Yes	Yes
R^2	0.681	0.654	0.623
Observations	92,820	94,910	92,820

Notes. The dependent variables are measured at the buyer-week level. *Value* is the sum of the dollar value of all completed contracts posted by a buyer in a given week, converted from local currency using contemporaneous exchange rates. *Contracts* is the number of completed contracts posted by a buyer in a given week. *Mean Value* is the average value of completed contracts. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a buyer-level exposure measure defined as the maximum automation exposure to generative AI across all contracts posted by the buyer prior to the introduction of ChatGPT and is normalized to the unit interval. All specifications include buyer fixed effects and source country-by-week fixed effects. Standard errors, clustered at the buyer level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Changes in Buyers' Willingness to Pay

	(1) Log(Max Budget)	(2) Log(Min Budget)	(3) Log(Budget Spread)
Post-GPT × Exposure	0.789***	0.564***	0.980***
	(0.203)	(0.208)	(0.303)
Constant	4.034***	3.315***	3.570***
	(0.078)	(0.079)	(0.120)
Buyer FE	Yes	Yes	Yes
Source Country × Time FE	Yes	Yes	Yes
R^2	0.608	0.644	0.632
Observations	94,657	94,779	48,164

Notes. The dependent variables are measured at the buyer-week level. *Max Budget* and *Min Budget* are the average maximum and minimum budgets of contracts posted by a buyer in a given week, converted from local currency using contemporaneous exchange rates. *Budget Spread* is the average difference between maximum and minimum budgets in contracts posted by a buyer in a given week. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a buyer-level exposure measure defined as the maximum automation exposure to generative AI across all contracts posted by the buyer prior to the introduction of ChatGPT and is normalized to the unit interval. All specifications include buyer fixed effects and source country-by-week fixed effects. Standard errors, clustered at the buyer level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Task Reorganization: Project Complexity and Interactivity

	(1) Log(Turnover Time)	(2) Log(Description)	(3) AI Share	(4) Interactive Share
Post-GPT × Exposure	1.196***	0.393**	0.023**	0.090
	(0.301)	(0.190)	(0.009)	(0.080)
Constant	1.597***	3.246***	−0.006	0.241***
	(0.115)	(0.073)	(0.004)	(0.030)
Buyer FE	Yes	Yes	Yes	Yes
Source Country × Week FE	Yes	Yes	Yes	Yes
R^2	0.521	0.519	0.331	0.425
Observations	70,804	94,910	94,910	94,910

Notes: The dependent variables are measured at the buyer-week. *Turnover Time* is the number of days between project posting and the date a review is left. *Description Length* is measured as the number of words in the project description. *AI Share* is the fraction of projects in a buyer-week that reference large language models or LLM-related tasks. *Interactive Share* is the fraction of projects in a buyer-week that require communication, interactions, and other interpersonal skills. *Post-GPT* is an indicator equal to one for weeks after November 30, 2022. *Exposure* is a buyer-level exposure measure defined as the maximum automation exposure to generative AI across all contracts posted by the buyer prior to the introduction of ChatGPT and is normalized to the unit interval. All specifications include buyer fixed effects and source country-by-week fixed effects. Standard errors are clustered at the buyer level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: AI Intensity and Outsourcing Margins

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Mean Value)
Post-GPT × AI Intensity	−0.481***	0.060***	−0.527***
	(0.140)	(0.019)	(0.138)
Constant	4.654***	0.079***	4.569***
	(0.001)	(0.000)	(0.001)
Buyer FE	Yes	Yes	Yes
Source Country × Week FE	Yes	Yes	Yes
R^2	0.581	0.354	0.593
Observations	92,820	94,910	92,820

Notes. The dependent variables are measured at the buyer-week level. *Value* is the sum of the dollar value of all completed contracts posted by a buyer in a given week, converted from local currency using contemporaneous exchange rates. *Contracts* is the number of completed contracts posted by a buyer in a given week. *Mean Value* is the average value of completed contracts. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *AI Intensity* is the share of projects involving AI-related tasks sourced by the buyer prior to ChatGPT's release and is normalized to the unit interval. All specifications include buyer fixed effects and source country-by-week fixed effects. Standard errors, clustered at the buyer level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Worker-Level Contract Adjustments

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Mean Value)
Post-GPT × Exposure	-0.190	-0.905***	0.663***
	(0.277)	(0.130)	(0.243)
Constant	5.185***	0.803***	4.405***
	(0.141)	(0.066)	(0.124)
Individual FE	Yes	Yes	Yes
Destination Country × Time FE	Yes	Yes	Yes
R^2	0.432	0.445	0.448
Observations	128,102	131,061	128,102

Notes. The dependent variables are measured at the worker-week level for continuing workers. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the number of completed contracts in a given week. *Mean Value* is the average value of completed contracts. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a worker-level exposure measure defined as the maximum exposure to GPT across all contracts performed by the worker prior to the introduction of ChatGPT and is normalized to the unit interval. All specifications include individual FE and destination country-by-week FE. Standard errors, clustered at the individual level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Worker-Level Skill Reallocations

	Skill Portfolio			Task Composition		
	Total	New	Existing	Skills/Contract	AI Share	Interact. Share
	(1)	(2)	(3)	(4)	(5)	(6)
Post-GPT × Exposure	−3.810***	0.901**	−4.711***	0.636***	0.122**	0.354***
	(0.724)	(0.370)	(0.666)	(0.187)	(0.051)	(0.090)
Constant	7.606***	0.720***	6.886***	4.262***	0.196***	0.160***
	(0.369)	(0.188)	(0.339)	(0.095)	(0.026)	(0.046)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Dest. Country × Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.338	0.390	0.491	0.215	0.599	0.207
Observations	131,061	131,061	131,061	131,061	131,061	131,061

Notes. The unit of observation is the worker–week for continuing workers. *Skill Count* is the number of distinct skills in which a worker completes at least one contract in a given week. *New Skills* counts skills used for the first time in a given week. *Existing Skills* counts skills previously used by the worker. *Skills/Contract* is the average number of skills associated with completed contracts. *AI Share* is the fraction of projects explicitly referencing large language models or AI-related tasks. *Interactive Share* is the fraction of projects requiring communication, coordination, creativity, or other interpersonal skills. *Post-GPT* equals one for weeks after November 30, 2022. *Exposure* is a worker-level pre-GPT exposure to generative AI, normalized to $[0, 1]$. All specifications include individual fixed effects, time fixed effects, and country-by-week fixed effects. Standard errors clustered at the individual level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Reskilling: Heterogeneity by Experience and Rating

	(1) Add Skill	(2) Add AI Skill	(3) Add Int. Skill	(4) Drop Skill
Panel A: Heterogeneity by Experience				
Post-GPT × Exposure	-0.407^{***}	-0.364^{***}	-0.330^{***}	0.867^{***}
	(0.102)	(0.055)	(0.110)	(0.096)
× ln(Pre-GPT Projects)	0.022^{***}	0.031^{***}	-0.009	-0.098^{***}
	(0.004)	(0.002)	(0.005)	(0.004)
Individual FE	Yes	Yes	Yes	Yes
Destination Country × Time FE	Yes	Yes	Yes	Yes
R^2	0.273	0.174	0.235	0.293
Observations	130,595	130,595	130,595	130,595
Panel B: Heterogeneity by Rating				
Post-GPT × Exposure	-0.070	0.174^{**}	-0.584^{***}	-0.139
	(0.105)	(0.070)	(0.147)	(0.117)
× Pre-GPT Rating	0.001	-0.009	0.023^{*}	-0.107^{***}
	(0.011)	(0.010)	(0.013)	(0.011)
Individual FE	Yes	Yes	Yes	Yes
Destination Country × Time FE	Yes	Yes	Yes	Yes
R^2	0.273	0.177	0.236	0.301
Observations	130,595	130,595	130,595	130,595

Notes. The unit of observation is the worker–week. *Add Skill* equals one if a worker performs at least one task in a new skill not previously supplied. *Add AI Skill* and *Add Int. Skill* indicate additions of AI-related and interaction-intensive skills, respectively. *Drop Skill* equals one if a worker ceases to supply at least one previously used skill. *Post-GPT* equals one for weeks after November 30, 2022. *Exposure* is a worker-level pre-GPT exposure to generative AI, normalized to $[0, 1]$. Standard errors clustered at the individual level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Earnings Impact of Reskilling by Skill Type

	Dependent Variable: Log(Value)				
	(1)	(2)	(3)	(4)	(5)
Post-GPT × Exposure	−0.261	−0.106	−0.642	0.136	0.005
	(0.273)	(0.426)	(1.410)	(0.280)	(0.274)
× Skill Additions	0.374***				
	(0.044)				
× AI Skill Additions		0.934***			
		(0.302)			
× Interactive Skill Additions			0.498**		
			(0.217)		
× Skill Reductions				−0.302***	
				(0.045)	
× Switching Intensity					0.215***
					(0.026)
Individual FE	Yes	Yes	Yes	Yes	Yes
Destination Country × Time FE	Yes	Yes	Yes	Yes	Yes
R^2	0.428	0.422	0.428	0.427	0.428
Observations	124,253	92,193	71,643	124,253	124,253

Notes: The dependent variable *Log(Value)* is the log total dollar value of completed contracts per worker-week, converted from local currency using contemporaneous exchange rates. *Post-GPT* is an indicator equal to one for weeks after November 30, 2022. *Exposure* is a worker-level measure of pre-GPT exposure to generative AI, normalized to $[0, 1]$. *Skill Additions* counts newly added skills. *AI Skill Additions* counts newly added AI-related skills (e.g., machine learning, AI integration, chatbot development). *Interactive Skill Additions* counts newly added interaction-related skills, such as those involving communication or client interaction. *Skill Reductions* measures reductions in pre-GPT skills. *Switching Intensity* measures the net change in the number of active skills relative to the pre-GPT period. All skill measures are normalized by the average number of skills in the pre- and post-GPT periods. All specifications include individual fixed effects and destination country-by-week fixed effects. Standard errors clustered at the individual level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Concentration Decomposition: Summary of Mechanisms

	Earnings	Projects
A. Does concentration rise in high-exposure skills?		
Within-skill concentration	0.574***	0.545***
	(0.061)	(0.061)
B. Does concentration shift across skills?		
Cross-skill contribution	−0.062	−0.085**
	(0.042)	(0.041)
C. What drives within-skill concentration?		
Participation (exit/entry)	−0.236***	−0.476***
	(0.056)	(0.071)
Within-active (stayers)	0.143***	0.348***
	(0.018)	(0.042)
D. Implied aggregate effect		
Within-skill ($\beta \times \bar{E}$)	0.275	0.261
Cross-skill ($\beta \times \bar{E}$)	−0.03	−0.04
Total concentration	**0.245**	**0.22**

Notes: Each entry reports the coefficient on skill-level generative AI exposure interacted with a post-GPT indicator from a skill–week regression. Panel A reports estimates from regressions in which the dependent variable is within-skill concentration, measured as the top-five share of earnings or projects within a skill–week. Panel B reports estimates from regressions where the dependent variable is the cross-skill component of aggregate concentration, constructed by weighting skill-level concentration measures by pre-GPT earnings shares. Panel C reports estimates from regressions that further decompose within-skill concentration into a participation component, capturing changes driven by entry and exit of workers within skills, and a within-active component, capturing changes in the distribution of outcomes among continuing workers. Panel D reports the implied contribution of within- and cross-skill channels to aggregate concentration. All specifications include skill and week fixed effects. Standard errors, reported in parentheses, are clustered at the skill level. The participation and within-active estimates in Panel C are based on different subsamples and therefore do not mechanically sum to the within-skill estimate in Panel A. Significance levels: *** $p < 0.01$, ** $p < 0.05$.

A A Theoretical Framework for AI-Driven Task Reallocation

This appendix develops a simple, stylized theoretical framework to guide the empirical analysis and illustrate how generative AI may reshape firms' outsourcing decisions, alter contract structure, induce worker reallocation, and generate concentration of market opportunities. The objective is not to deliver a fully structural account of organizational choice, but to make transparent how the technological change maps into the estimable objects in our data.

A.1 Setup

A.1.1 Tasks

A representative firm must complete a unit measure of tasks indexed by $i \in [0, 1]$ to produce output. Tasks are characterized by two attributes:

- **Automation exposure** $a_i \in [0, 1]$: the degree to which generative AI can autonomously perform task i, reducing the need for human labor input.
- **Complexity** $b_i \in [0, 1]$: the degree to which task i requires coordination across inputs, client-specific customization, iterative feedback, or exercise of judgment.

We assume that exposure and complexity are negatively correlated in the task distribution:

$$\mathbb{E}[b_i \mid a_i] = \bar{b} - \rho a_i, \quad \rho > 0. \tag{A-1}$$

This structure reflects the empirical regularity that highly automatable tasks—routine coding, translation, standardized content generation—tend to be modular and codifiable, while tasks that resist automation—strategic advising, system integration, creative direction—tend to require tacit knowledge and adaptive judgment.[32]

Let $H(a, b)$ denote the joint distribution of tasks over exposure and complexity.

A.1.2 Production Modes

Each task can be completed through one of three modes:

1. **Automation**: performed using generative AI at unit cost $\kappa_i(\theta)$.

[32] This negative correlation is consistent with the task-based literature distinguishing routine from non-routine work (Autor et al., 2003; Acemoglu and Autor, 2011). In our setting, automation exposure captures susceptibility to generative AI specifically, while complexity captures the coordination and customization requirements that make tasks difficult to fully specify in arm's-length contracts.

2. **Outsourcing**: performed by an external contractor at unit cost c_{ik} depending on contractor location k.

3. **Internal production**: performed in-house at reservation cost $\bar{w}$.

We focus on the automation–outsourcing margin, treating internal production as an outside option that absorbs tasks for which neither automation nor outsourcing is cost-effective. This focus is appropriate for our empirical setting, which observes realized outsourcing transactions on a platform marketplace.

A.1.3 Automation Technology

If task i is performed using generative AI, the unit cost is

$$\kappa_i(\theta) = \bar{\kappa} - \theta a_i, \tag{A-2}$$

where $\theta \geq 0$ indexes the capability of frontier AI systems. The parameter $\bar{\kappa}$ captures baseline costs of deploying AI (infrastructure, integration, prompt engineering, quality assurance), and θa_i captures capability-driven cost reductions that are increasing in task exposure.

Improvements in AI capability correspond to increases in θ. This formulation implies that capability improvements reduce automation costs for all tasks but do so most strongly for tasks with high exposure a_i. The discrete releases of GPT-3.5, GPT-4, and GPT-4o in our sample period can be interpreted as increases in θ that shift the automation cost frontier.

A.1.4 Outsourcing Technology

Outsourced tasks are performed by contractors located in either developing economies (f) or developed economies (d). The unit cost of outsourcing task i to destination $k \in \{f, d\}$ is

$$c_{ik} = \frac{w_k}{\phi_k(b_i)}, \tag{A-3}$$

where w_k is the destination wage and $\phi_k(b_i)$ is effective productivity, which depends on task complexity.

Wages. We assume $w_f < w_d$: developing-country contractors have lower wages, reflecting differences in outside options, cost of living, and labor market conditions. This wage differential is the fundamental driver of outsourcing to developing countries and is well documented in the literature on offshore services trade.

Productivity. We assume that productivity is similar across destinations for routine, well-specified tasks, but that developed-country contractors have a productivity advantage in complex tasks requiring judgment, coordination, or client-specific adaptation:

$$\phi_f(b) = \phi_0, \qquad \phi_d(b) = \phi_0 + \mu b, \tag{A-4}$$

where $\phi_0 > 0$ is baseline productivity common across destinations and $\mu > 0$ captures the developed-country productivity premium in complex tasks.

This structure reflects that routine, codifiable tasks such as data entry, basic translation, and standardized coding can be performed comparably well by workers in either location. The task requirements can be fully specified and transmitted digitally, and execution does not depend on physical proximity, specialized institutional knowledge, or intensive client interaction. Complex tasks, in contrast, benefit from the specialized skills, established client relationships, and institutional context more prevalent among developed-country workers.

Comparative advantage. Under these assumptions, the relative cost of developing- versus developed-country outsourcing is

$$\frac{c_{if}}{c_{id}} = \frac{w_f}{w_d} \cdot \frac{\phi_0 + \mu b_i}{\phi_0}. \tag{A-5}$$

For low-complexity tasks (b_i small), this ratio is approximately $w_f/w_d < 1$: developing countries are cheaper because the wage advantage dominates and there is no offsetting productivity differential. For high-complexity tasks (b_i large), the ratio increases and may exceed one if the developed-country productivity premium is sufficiently large relative to the wage gap.

Developing countries thus have comparative advantage in low-complexity tasks, where the wage differential determines relative costs. Developed countries have comparative advantage in high-complexity tasks, where productivity differences can offset higher wages. Combined with the negative correlation between exposure and complexity in (A-1), this implies that high-exposure tasks are disproportionately outsourced to developing countries, a pattern documented in the data.

A.1.5 Contractors

Within each destination k, contractors are heterogeneous in ability η, distributed according to $G_k(\eta)$ with support $[\underline{\eta}, \bar{\eta}]$, where $0 < \underline{\eta} < \bar{\eta}$. Ability affects productivity: a contractor with ability η performing task i in destination k has effective productivity $\eta \cdot \phi_k(b_i)$.

We assume ability and complexity are (weakly) complementary:

$$\frac{\partial^2(\eta \cdot \phi_k)}{\partial \eta \, \partial b} \geq 0, \quad \text{with strict inequality when } \frac{\partial \phi_k}{\partial b} > 0. \tag{A-6}$$

Under equation (A-4), this cross-partial equals $\mu > 0$ for developed-country contractors ($k = d$), since $\partial^2[\eta(\phi_0 + \mu b)]/\partial \eta \, \partial b = \mu$, and zero for developing-country contractors ($k = f$), whose productivity is constant in complexity. The operative content of assumption (A-6) is therefore that, among developed-country contractors, high-ability workers have comparative advantage in complex tasks that reward skill, judgment, and adaptability. For developing-country contractors, ability raises productivity uniformly across task types but does not interact with complexity, consistent with their specialization in routine work where the scope for differentiation is limited.[33]

A.1.6 Contract Structure

Outsourced tasks are bundled into contracts. A contract specifies a set of tasks $S \subseteq [0, 1]$ to be performed by a single contractor. Contracting is costly: each contract incurs a fixed coordination cost $F > 0$ that captures search, matching, contract specification, monitoring, and payment processing.

The total cost of a contract covering task set S assigned to a contractor in destination k is

$$C(S; k) = F + \int_{i \in S} c_{ik} \, di. \tag{A-7}$$

The fixed cost F creates an incentive to bundle tasks into contracts rather than procuring each task separately. However, the optimal scope of bundling depends on the characteristics of the tasks being outsourced.

Routine tasks can be precisely specified, easily verified, and executed independently, favoring separate procurement with specialized contractors. Complex tasks, in contrast, favor consolidation: they are interdependent (requiring coordination), incompletely specified (requiring adaptation), and benefit from accumulated client-specific knowledge. As outsourced work shifts toward complex tasks, optimal contract structure shifts toward fewer, broader engagements.

[33] The zero cross-partial for developing-country contractors is a consequence of our simplifying assumption that ϕ_f is independent of b. If we instead assumed a small positive slope $\phi_f(b) = \phi_0 + \mu_f b$ with $0 < \mu_f \ll \mu$, the complementarity would hold strictly for both destinations without altering any qualitative results. We maintain the simpler specification for tractability.

A.2 Buyer Task Allocation and Consolidation

A.2.1 Mode Choice

The firm assigns each task to its least-cost production mode. For task i, the firm compares:

$$\min\left\{\kappa_i(\theta),\ \min_{k\in\{f,d\}} c_{ik},\ \bar{w}\right\}. \tag{A-8}$$

Define the *automation threshold* $\hat{a}(\theta)$ as the exposure level at which automation and the best outsourcing option have equal cost:

$$\bar{\kappa} - \theta\hat{a} = \min_k c_k(\mathbb{E}[b \mid a = \hat{a}]). \tag{A-9}$$

Tasks with exposure $a_i > \hat{a}(\theta)$ are automated; tasks with $a_i < \hat{a}(\theta)$ are outsourced (provided outsourcing is cheaper than internal production) or produced internally. The threshold $\hat{a}(\theta)$ is the margin between automation and outsourced labor in the task space.

A.2.2 Comparative Statics

Proposition 1 (Effect of AI capability on task allocation)**.** *An increase in AI capability θ has the following effects:*

1. *The automation threshold falls:* $\dfrac{\partial\hat{a}}{\partial\theta} < 0$.
2. *The measure of outsourced tasks contracts:* $\dfrac{\partial|\mathcal{I}_O|}{\partial\theta} < 0$.
3. *The average exposure of outsourced tasks falls:* $\dfrac{\partial}{\partial\theta}\mathbb{E}[a_i \mid i \in \mathcal{I}_O] < 0$.
4. *The average complexity of outsourced tasks rises:* $\dfrac{\partial}{\partial\theta}\mathbb{E}[b_i \mid i \in \mathcal{I}_O] > 0$.

Proof. Part (1): From (A-9), an increase in θ lowers the left-hand side (reduces automation cost) for any fixed $\hat{a}$. To restore equality, $\hat{a}$ must fall, which raises $\bar{\kappa} - \theta\hat{a}$ back toward the outsourcing cost. Hence $\partial\hat{a}/\partial\theta < 0$.

Part (2): The outsourced set is $\mathcal{I}_O(\theta) = \{i : a_i < \hat{a}(\theta) \text{ and } \min_k c_{ik} < \bar{w}\}$. Since $\hat{a}$ falls with θ, the set of tasks satisfying $a_i < \hat{a}(\theta)$ shrinks. Hence $|\mathcal{I}_O|$ contracts.

Parts (3) and (4): The tasks removed from $\mathcal{I}_O$ as θ increases are those with exposure just below the old threshold—i.e., the highest-exposure tasks in the outsourced set. By the negative correlation in (A-1), these are also the lowest-complexity tasks on average. Removing high-exposure, low-complexity tasks from $\mathcal{I}_O$ lowers the mean exposure and raises the mean complexity of the remaining outsourced tasks. ∎

A.2.3 Destination-Specific Effects

Proposition 2 (Asymmetric displacement across destinations)**.** *Displacement of outsourcing is concentrated in developing countries. Formally, let* $\mathcal{I}_f(\theta)$ *and* $\mathcal{I}_d(\theta)$ *denote the sets of tasks outsourced to developing and developed countries, respectively. Then:*

$$\left|\frac{\partial|\mathcal{I}_f|}{\partial\theta}\right| > \left|\frac{\partial|\mathcal{I}_d|}{\partial\theta}\right|.$$

Proof. By equation (A-1), high-exposure tasks have low complexity on average. By equation (A-5), developing countries have lower costs for low-complexity tasks because their wage advantage dominates when the developed-country productivity premium μb is small. Thus, prior to the increase in θ, high-exposure tasks are disproportionately outsourced to developing countries: $\mathcal{I}_f$ is concentrated in the high-a, low-b region of the task space.

The tasks displaced by automation as θ increases are those with exposure near the threshold $\hat{a}(\theta)$, i.e., the highest-exposure tasks remaining in the outsourced set. These tasks are disproportionately drawn from $\mathcal{I}_f$. Hence, the contraction in outsourcing falls more heavily on developing-country destinations. Developed-country outsourcing $\mathcal{I}_d$ is concentrated in low-exposure, high-complexity tasks where the productivity premium justifies higher wages. These tasks remain well below the automation threshold as θ increases, so $|\mathcal{I}_d|$ is relatively unaffected. ■

The key mechanism is straightforward: developing countries' wage advantage gives them comparative advantage in routine tasks; routine tasks (high exposure, low complexity) are automated first as AI capability improves; therefore, developing countries bear the brunt of displacement.

A.2.4 Contract Consolidation

Proposition 3 (Consolidation)**.** *As AI capability increases, buyers consolidate remaining work into fewer, larger contracts.*

Proof. By Proposition 1, outsourced tasks shift toward higher complexity. Complex tasks are interdependent, harder to specify, and benefit from accumulated client knowledge, all of which favor relational contracting with a single contractor over separate procurement. Combined with fixed costs per contract, this implies consolidation. Higher complexity also implies higher contract values and more detailed specifications. ■

A.3 Worker Adjustment and Skilling

The preceding analysis characterizes firm-side responses to improvements in AI capability. We now turn to the supply side: how do workers adjust their skill portfolios in response to the changing composition of outsourced demand?

Notation. The analysis shifts from task-level to skill-level. In the preceding sections, i indexed individual tasks in $[0, 1]$ and k indexed contractor destinations. In the remainder of this section, we re-index: j denotes a worker's origin skill, i denotes a (potential) destination skill, and the set of all skills is $\mathcal{I}$. Each skill $i \in \mathcal{I}$ corresponds to a subset of tasks $\mathcal{T}_i \subset [0, 1]$, linking the skill-level analysis back to the task-level framework above.

A.3.1 Skills and Skill Value

Workers supply skills to the outsourcing market. Let $\mathcal{I}$ denote the set of skills available on the platform, where each skill $i \in \mathcal{I}$ corresponds to a subset of tasks $\mathcal{T}_i \subset [0, 1]$. Skills may overlap: a task may require multiple skills, and a worker may supply multiple skills.

Let $V_i(\theta)$ denote the expected value to a contractor of supplying skill i when AI capability is θ. This reduced-form object summarizes the equilibrium return to skill i, incorporating expected demand for the skill, expected compensation conditional on matching, and any continuation or reputation value. We do not model the determination of V_i from primitives but instead characterize how it responds to changes in θ through the demand-side mechanisms already established.

Lemma 1 (Skill values respond to AI capability)**.** *Improvements in AI capability affect skill values through two channels:*

1. ***Automation pressure****: Skills tied to high-exposure tasks experience declining demand and hence falling value as those tasks shift from outsourcing to automation.*

2. ***Reorganization complementarity****: Skills that facilitate coordination, integration, client interaction, or project management experience rising value as buyers consolidate into complex, multi-dimensional projects.*

Skills in the second category, which we term *AI-complementary*, gain value as the reorganization of outsourced work toward complex projects raises the demand for skills that support such projects. This is a composition effect operating through the demand side, distinct from direct productivity augmentation, which we explore in Section A.5.

A.3.2 Reskilling Technology

Workers can reallocate across skills by paying switching costs.[34] A worker currently supplying skill j who switches to a different skill $i \neq j$ incurs cost

$$c_{ji} = \begin{cases} 0 & \text{if } i = j \text{ (no switching)}, \\ c_0 + \tau \cdot d(j,i) & \text{if } i \neq j, \end{cases} \tag{A-10}$$

where $c_0 > 0$ is a fixed component reflecting the baseline cost of acquiring any new skill, $\tau > 0$ governs the importance of skill distance, and $d(j,i) \geq 0$ measures the technological distance between skills j and i. Remaining in the current skill incurs no adjustment cost.

Distance is symmetric ($d(j,i) = d(i,j)$) and satisfies the triangle inequality. We interpret $d(j,i)$ as reflecting the transferability of human capital: skills that share common competencies, tools, or domain knowledge have low distance, while skills requiring fundamentally different expertise have high distance. In the empirical analysis, we proxy for distance using pre-AI patterns of skill co-occurrence among workers and in project requirements.

A.3.3 Skill Choice

Workers choose which skill to supply to maximize expected utility. A worker currently supplying skill j evaluates each potential destination skill $i \in \mathcal{I}$ (including the option of remaining at j) according to

$$U_{ji} = V_i(\theta) - c_{ji} + \varepsilon_{ji}, \tag{A-11}$$

where ε_{ji} is an idiosyncratic taste shock. Assume $\{\varepsilon_{ji}\}_{i \in \mathcal{I}}$ are i.i.d. Type-I extreme value with scale parameter normalized to one.

Let M_j denote the mass of workers initially supplying skill j. Under the distributional assumption, the probability that a worker from skill j transitions to skill i is

$$P_{ji} = \frac{\exp\big(V_i(\theta) - c_{ji}\big)}{\sum_{\ell \in \mathcal{I}} \exp\big(V_\ell(\theta) - c_{j\ell}\big)}, \tag{A-12}$$

and the expected flow of workers from j to i is $\omega_{ji} = M_j \cdot P_{ji}$.

[34] We use c_{ji} for switching costs in this section. Context distinguishes this from the outsourcing cost c_{ik} in Sections A.1–A.2, where the first subscript indexes tasks and the second indexes destinations. Here, both subscripts index skills.

A.3.4 Gravity Structure

For transitions to new skills ($i \neq j$), taking logs of the flow equation and expanding:

$$\log \omega_{ji} = \log M_j + V_i(\theta) - c_0 - \tau d(j,i) - \log \Phi_j(\theta), \tag{A-13}$$

where $\Phi_j(\theta) = \sum_{\ell \in \mathcal{I}} \exp\big(V_\ell(\theta) - c_{j\ell}\big)$ is the inclusive value from origin skill j, incorporating both the option of staying ($\ell = j$, with $c_{jj} = 0$) and all switching alternatives.

With origin fixed effects α_j absorbing $\log M_j$, c_0, and $\log \Phi_j(\theta)$, we obtain the gravity specification for $i \neq j$:

$$\log \omega_{ji} = \alpha_j + V_i(\theta) - \tau \cdot d(j,i) + \nu_{ji}, \tag{A-14}$$

where ν_{ji} is a residual with conditional mean zero.

This structure is analogous to gravity equations in trade, with origin fixed effects, destination attractiveness (V_i), and bilateral frictions ($\tau \cdot d(j,i)$). It yields a directly estimable specification linking worker flows to skill proximity and destination characteristics.

Proposition 4 (Directed reskilling)**.** *Worker transitions across skills satisfy two properties:*

1. ***Proximity****: Conditional on destination value, flows are decreasing in skill distance:* $\partial \log \omega_{ji} / \partial d(j,i) = -\tau < 0$.

2. ***Direction****: Conditional on distance, flows are increasing in destination skill value:* $\partial \log \omega_{ji} / \partial V_i = 1 > 0$.

As AI capability increases, flows are directed toward skills whose value rises, particularly AI-complementary skills that benefit from contract consolidation and task reorganization.

The proximity property implies that reskilling is localized: workers move along nearby skill dimensions rather than making large jumps across unrelated domains. The direction property implies that, conditional on feasibility, workers move toward skills offering higher returns. These properties predict concentrated flows toward AI-complementary skills that are proximate to displaced skills.

Importantly, directed reskilling toward high-value destinations does not imply that workers flee automation exposure per se. Workers may move toward nominally high-exposure skill categories if those skills involve tasks that complement rather than compete with AI, such as deploying, customizing, supervising, or integrating AI systems. What matters for the direction of flows is the post-AI value $V_i(\theta)$, not the exposure level alone.

A.3.5 Heterogeneous Adaptation

Workers differ in their capacity to adjust. Let worker ability η affect the cost of switching to a new skill:

$$c_{ji}(\eta) = \begin{cases} 0 & \text{if } i = j, \\ c_0 + \tau \cdot d(j,i) - \gamma \cdot \eta & \text{if } i \neq j, \end{cases} \tag{A-15}$$

where $\gamma > 0$ implies that high-ability workers face lower switching costs. This could reflect greater learning capacity, more transferable general skills, better access to information about emerging opportunities, or financial resources to invest in retraining. Crucially, remaining at the current skill continues to incur zero cost regardless of ability: the asymmetry between staying and switching is what generates ability-dependent reallocation.

Corollary 1 (Heterogeneous adaptation)**.** *High-ability workers are more likely to reskill in response to changes in skill values induced by AI. Among workers initially supplying skill j, the probability of transitioning to any destination $i \neq j$ is increasing in ability η.*

Proof. Let $A = \exp(V_j(\theta))$ denote the attractiveness of staying, $B_i = \exp(V_i(\theta) - c_0 - \tau d(j,i))$ the base attractiveness of destination $i \neq j$, $S = \sum_{i' \neq j} B_{i'}$, and $\psi = \exp(\gamma\eta)$. Since staying incurs zero cost while switching incurs cost $c_0 + \tau d(j,i) - \gamma\eta$, the transition probability for $i \neq j$ is:

$$P_{ji}(\eta) = \frac{\psi\, B_i}{A + \psi\, S}.$$

Differentiating with respect to η:

$$\frac{\partial P_{ji}}{\partial \eta} = \gamma\psi \cdot \frac{A\, B_i}{\big(A + \psi\, S\big)^2} > 0,$$

since $A, B_i, \gamma, \psi > 0$. Hence $P_{ji}(\eta)$ is strictly increasing in η for every destination $i \neq j$: higher-ability workers are more likely to switch to any given destination. Since this holds for all $i \neq j$, the total probability of switching, $1 - P_{jj}(\eta) = \psi S/(A + \psi S)$, is also increasing in η.[35] ∎

[35] Conditional on switching, the allocation across destinations is $P_{ji}/(1 - P_{jj}) = B_i/S$, which is independent of η. High-ability workers are thus more likely to switch but do not differentially favor high-value destinations over low-value ones, conditional on switching. However, because more high-ability workers switch overall, they constitute a larger share of the transition flow to *every* destination $i \neq j$ than their share in the origin population.

A.4 Concentration

The above framework further implies that AI-driven reorganization increases concentration of jobs and earnings among contractors.

Proposition 5 (Increasing Concentration)**.** *An increase in AI capability raises concentration through three channels:*

1. ***Selection****: Demand contraction induces exit among contractors in high-exposure tasks. Among developed-country contractors, where ability and complexity are complementary by* (A-6)*, exiting contractors are disproportionately low-ability. Among developing-country contractors, exit reflects uniform demand contraction rather than ability-specific selection on complexity.*

2. ***Threshold effects****: Consolidation raises capability requirements; buyers bundling complex projects prefer contractors who can deliver across multiple dimensions, rationing out lower-ability contractors.*

3. ***Differential adaptation****: High-ability contractors reskill more effectively toward AI-complementary skills (Corollary 1), capturing disproportionate shares of growing categories.*

Together, these forces shift the distribution of jobs and earnings toward high-ability contractors, increasing concentration at the top of the distribution.

A.4.1 Decomposing Concentration

Changes in aggregate concentration can be decomposed into cross-skill and within-skill components. Let s_i denote skill i's share of platform earnings, and let C_i denote concentration (e.g., the top-5 share) within skill i. Aggregate concentration is $\bar{C} = \sum_i s_i C_i$, and its change decomposes as:

$$\Delta\bar{C} = \underbrace{\sum_i \bar{s}_i \Delta C_i}_{\text{within-skill}} + \underbrace{\sum_i \bar{C}_i \Delta s_i}_{\text{cross-skill}} + \text{interaction}, \tag{A-16}$$

where bars denote baseline (pre-GPT) values. The within-skill component captures changes in the distribution of earnings among contractors supplying the same skill. The cross-skill component captures shifts in platform activity across skill categories, for instance, if activity reallocates toward skills that were already highly concentrated.

The framework predicts that within-skill concentration rises in high-exposure skills through the selection and threshold mechanisms described above. The cross-skill component is ambiguous *a priori*: if high-exposure skills contract and these skills happened to be more

concentrated than average, the cross-skill component would be negative (activity shifts away from concentrated skills); the reverse would generate a positive cross-skill contribution.

The within-skill component can be further decomposed to distinguish the role of worker exit from rising inequality among continuing workers. Let C_i^{pre} denote concentration in skill i computed only among workers who were active before AI introduction. Then:

$$\Delta C_i = \underbrace{(C_i - C_i^{\text{pre}})}_{\text{participation}} + \underbrace{(C_i^{\text{pre}} - \bar{C}_i)}_{\text{within-active}}. \tag{A-17}$$

The *participation* margin captures how worker entry and exit affect measured concentration. The *within-active* margin captures changes in concentration among workers who were active before AI and remain active afterward. This component isolates the pure selection and threshold effects operating within a fixed set of workers: as capability requirements rise and select workers reskill, top performers may capture a growing share, even without any change in the composition of the workforce.

The framework thus generates three testable predictions regarding concentration:

1. *Within-skill concentration rises*: High-exposure skills experience greater increases in concentration than low-exposure skills.

2. *Cross-skill reallocation is ambiguous*: The sign depends on whether high-exposure skills were more or less concentrated than average prior to AI introduction.

3. *Within-active concentration rises*: Among workers who remain active, top performers capture a growing share.

The participation margin provides a diagnostic for the nature of selection: a positive coefficient indicates that exit is concentrated among low performers, while a negative coefficient indicates that some top performers also exit, potentially due to heterogeneity in adjustment costs or outside options not captured in the framework.

A.5 Extensions

We consider two extensions that address additional empirical patterns and potential alternative mechanisms.

A.5.1 AI-Assisted Production and Quality

The baseline model treats automation as full substitution: tasks are either performed by AI or by human contractors, not both. In practice, generative AI may also augment human

productivity, enabling workers to complete tasks faster or produce higher-quality output.

To capture this possibility, let output quality for a contractor with ability η in destination k performing task i be:

$$q_i(\eta; \theta, k) = \left[\eta + \delta(\theta) \cdot a_i\right] \cdot \phi_k(b_i), \tag{A-18}$$

where $\delta(\theta) \geq 0$, with $\delta'(\theta) > 0$, captures AI-assisted quality improvements. The additive term $\delta(\theta) \cdot a_i$ represents a capability boost from AI tools that is independent of the contractor's baseline ability and increasing in both AI capability and task exposure: AI tools provide greater assistance for tasks where AI is more capable.[36]

This extension generates additional predictions:

Quality improvements among continuing workers. Among contractors who remain active on the platform, AI assistance raises output quality. The magnitude of improvement is larger for (i) high-exposure tasks, where AI tools are most effective, and (ii) lower-ability workers, for whom the additive boost $\delta(\theta) \cdot a_i$ represents a proportionally larger gain relative to baseline quality $\eta \cdot \phi_k(b_i)$.

Compression of quality differentials. Prior to AI, quality differences across contractors within a given destination were driven primarily by ability differences: $q_i(\eta; 0, k) = \eta \cdot \phi_k(b_i)$ (assuming $\delta(0) = 0$). With AI assistance, the gap between high- and low-ability contractors narrows for high-exposure tasks. For any destination k:

$$\frac{\partial}{\partial \theta}\left[\frac{q_i(\eta_H; \theta, k)}{q_i(\eta_L; \theta, k)}\right] = \frac{\delta'(\theta)\, a_i\, (\eta_L - \eta_H)}{\left[\eta_L + \delta(\theta)\, a_i\right]^2} < 0 \quad \text{for } a_i > 0, \tag{A-19}$$

since $\eta_L < \eta_H$ and $\delta'(\theta) > 0$. The inequality is strict whenever $a_i > 0$, and the compression is stronger for tasks with higher automation exposure. This "leveling" effect is consistent with evidence that generative AI compresses productivity distributions by disproportionately benefiting lower performers (Brynjolfsson et al., 2025; Noy and Zhang, 2023).

Interaction with displacement. The augmentation and substitution effects operate simultaneously but on different margins. Substitution removes high-exposure tasks from the outsourced set entirely. Augmentation raises quality for high-exposure tasks that remain outsourced (e.g., because full automation is not yet cost-effective or because human oversight is required). The net effect on any given task depends on whether it falls above or

[36] The additive specification is essential for generating quality compression across ability levels. A multiplicative specification $q_i(\eta; \theta, k) = \eta \cdot \phi_k(b_i) \cdot [1 + \delta(\theta) a_i]$ would scale quality proportionally for all workers, leaving the ratio $q_i(\eta_H; \theta, k)/q_i(\eta_L; \theta, k) = \eta_H/\eta_L$ independent of θ and generating no convergence.

below the automation threshold $\hat{a}(\theta)$.

A.5.2 Task Expansion

The baseline model holds the set of tasks fixed and analyzes how AI reallocates existing tasks across production modes. An alternative possibility is that AI creates new tasks or expands the scope of feasible projects.

To capture this, let the measure of tasks depend on AI capability:

$$|\mathcal{T}(\theta)| = \bar{T} + \zeta(\theta), \tag{A-20}$$

where $\bar{T} = 1$ is the baseline task measure (corresponding to the unit interval in the baseline model) and $\zeta(\theta) > 0$ is an expansion term increasing in θ. New tasks created by AI—such as prompt engineering, AI system integration, model fine-tuning, AI-assisted data analysis, and human-AI workflow design—are indexed by high values of i and have characteristics that differ systematically from pre-existing tasks.

Characteristics of new tasks. We assume that AI-created tasks have moderate-to-high automation exposure (they involve working with AI systems) but also high complexity (they require judgment, customization, and coordination):

$$\mathbb{E}[a_i \mid i \in \mathcal{T}(\theta) \setminus \mathcal{T}(0)] = a^{\text{new}} > \mathbb{E}[a_i], \qquad \mathbb{E}[b_i \mid i \in \mathcal{T}(\theta) \setminus \mathcal{T}(0)] = b^{\text{new}} > \mathbb{E}[b_i], \tag{A-21}$$

where expectations on the right-hand side are taken over the pre-AI task distribution H. These tasks are "AI-complementary" in the sense that they require human expertise to deploy, customize, and oversee AI systems rather than tasks that AI performs autonomously.

Implications. Task expansion partially offsets displacement: while AI automates some existing high-exposure tasks, it also creates new tasks that require human labor. The net effect on total outsourcing volume is ambiguous and depends on the relative magnitudes of $\partial\hat{a}/\partial\theta$ (displacement) and $\partial\zeta/\partial\theta$ (expansion).

However, the composition of outsourced work shifts unambiguously. New AI-complementary tasks are high-complexity by assumption, reinforcing the shift toward complex work documented in Proposition 1. Workers who can supply skills relevant to AI-complementary tasks, such as integration, customization, oversight, and prompt engineering, benefit from task expansion, while workers specialized in routine tasks that are automated experience displacement without offsetting demand from new task creation.

Task expansion also affects the geography of outsourcing. If AI-complementary tasks have high complexity (b^{new} large), they favor developed-country workers who have comparative advantage in complex work. This would further concentrate displacement in developing countries: not only do developing countries lose demand for routine tasks (baseline model), but the new tasks created by AI flow disproportionately to developed countries.

Empirical implications. Task expansion predicts: (i) growth in explicitly AI-related project categories (prompt engineering, AI integration, chatbot development), (ii) these projects command premium compensation reflecting high complexity, and (iii) demand for AI-related skills concentrates in developed countries or among high-ability contractors in developing countries who can meet the capability requirements. These predictions are testable using variation in AI-related keywords in project descriptions and skill tags.[37]

A.6 Summary of Predictions

The framework generates four predictions that structure the empirical analysis:

1. **Task displacement**: Generative AI displaces outsourcing in high-exposure tasks. The automation threshold shifts inward, removing routine cognitive tasks from the outsourced set. Contractors in developing countries whose comparative advantage lies in precisely these tasks bear the brunt of displacement.

2. **Contract consolidation**: Buyers consolidate remaining demand into fewer, higher-value contracts that bundle related tasks under single contractors capable of coordination and integrated delivery. The composition of outsourced work shifts toward complex, interaction-intensive engagements.

3. **Directed reskilling**: Workers reallocate toward technologically proximate skills with higher expected returns. High-ability workers adapt more effectively.

4. **Rising concentration**: Selection, rising capability thresholds, and differential adaptation increase the concentration of earnings and contracts among top performers within skill categories.

We examine these predictions empirically by exploiting variation in AI exposure across skills, buyers, and workers, and transaction-level data that allow us to observe realized contracts, task reallocation, and skill adjustment at high frequency.

[37] It is worth noting that while the scope of the framework is appropriate for our setting, the model abstracts from platform competition and the possibility that AI-driven reorganization could induce reallocation across platforms, not just within them.

B Data Appendix

This appendix provides detailed information on the data sources, collection procedures, sampling design, variable construction, and limitations of the dataset. The purpose is to ensure transparency and reproducibility and to clarify the scope and interpretation of the empirical results.

Data sources and collection. All data used in this study are collected from the API of the online labor marketplace analyzed in the paper. Data were retrieved using automated scripts that collect worker profiles, client profiles, project listings, and contract histories visible to any user browsing the platform. Collection procedures complied with the platform's rate limits and did not circumvent access controls or authentication requirements.

The dataset covers the period from January 2022 through December 2024. This window spans the introduction of generative AI tools and allows us to observe both pre-introduction outcomes and post-introduction adjustments in outsourcing activity.

The raw data consist of three linked units of observation: worker profiles, client profiles, and project-level contracts. These units are merged using platform-assigned identifiers. Prior to analysis, we perform a series of data-cleaning and consistency checks. These include removing duplicate observations arising from repeated page retrievals, verifying that project completion dates follow posting dates, standardizing currencies and timestamps, and dropping observations with missing identifiers or implausible values. These procedures ensure that the final dataset contains internally consistent and complete transaction histories.

Sampling design and cohort construction. A defining feature of online labor marketplaces is the extreme concentration of economic activity. Although the platform hosts a large number of registered users, only a small fraction ever completes a contract, and an even smaller fraction does so regularly. As a result, nearly all realized transactions are generated by a limited subset of active workers. For the purpose of studying outsourcing behavior and its response to technological change, comprehensive collection of all registered accounts is neither feasible nor informative.

To focus on the economically relevant portion of the market, we construct a fixed cohort of workers with engagement prior to the diffusion of generative AI. For each of the platform's 2,700 skill categories, we identify workers who completed at least one contract in that skill before November 2022 and retain up to 500 per skill following the platform's internal ranking system based on completed contracts and client ratings. The 500-worker threshold reflects the platform's matching architecture: with 10 workers displayed per page, reaching the 500th-ranked worker requires navigating 50 pages of search results, well beyond the depth

at which buyers typically engage (Ursu, 2018). Given the granularity of our skill definitions, over 95 percent of skill categories have fewer than 500 workers meeting our inclusion criterion; for these categories, our sample includes the universe of active participants. For the remaining larger categories, our sample captures the workers who appear in buyer searches and generate the observed transactions we study. Because workers frequently supply multiple skills (median: 8 categories), this procedure yields approximately 93,000 unique individuals. In the pre-period, these workers accounted for approximately 95 percent of all completed contracts on the platform, consistent with the extreme concentration characteristic of digital labor markets (Horton, 2010; Agrawal et al., 2016).

The cohort is fixed at entry: workers are included based on pre-November 2022 activity and followed throughout the sample period regardless of whether they remain active, exit the platform, or change their skill composition.[38] This design captures both intensive-margin adjustment among continuing workers and extensive-margin displacement through exit within a single cohort observed over the full sample period.

Client-worker-skill transaction data and variable definitions. For each worker in the cohort, we collect the complete history of posted and completed projects visible on the worker's profile between January 2022 and December 2024. The resulting dataset contains over 300,000 projects. About 91 percent of these projects are cross-border transactions in which the client and worker report different countries of residence, reflecting the platform's primary function as an intermediary for international outsourcing of digitally deliverable services. Worker characteristics are drawn from publicly visible profile information and include country of residence, tenure on the platform, education and certifications when available, self-reported skills, historical job counts, and client ratings. Client characteristics are extracted from project and client pages and include client country, industry category when available, hiring history, and ratings.

Project-level information includes the project title and description, associated skill tags, contract type (hourly or fixed-price), posting and completion dates, budget ranges, number of bids, hours worked for hourly contracts, realized payments when visible, and multidimensional client evaluations covering quality, communication, timeliness, and professionalism. All monetary values are converted to a common currency using contemporaneous exchange rates. From the project data, we construct outcomes at multiple levels to study different margins of adjustment. At the skill–week level, we compute total project counts, total contract value, and average contract value. At the client–week level, we aggregate the number

[38] Fixing the cohort ensures that changes in outsourcing volumes, earnings, or task composition are not mechanically driven by confounding shifts in platform composition and traffic.

and value of projects and the distribution of demand across skills. At the worker–week level, we measure the number of projects completed, total earnings, average project value, and the distribution of work across skills. These aggregations allow us to analyze adjustment along extensive, intensive, and compositional margins of activity in the market.

Table A.1 reports the largest source countries, destination countries, and professional categories observed in the dataset and confirms that the platform is primarily used for international outsourcing of digitally deliverable services. Outsourcing flows in the data originate predominantly from high-income countries; the United States accounts for 27.4% of projects, followed by Australia (10%), the United Kingdom (8.1%), and Canada (5.5%). The projects are directed mostly toward developing-country destinations; approximately 60% of contracts are performed in India and Pakistan, followed by Bangladesh and Egypt. The tasks are heavily concentrated in professional services such as design, media, IT, software development, content creation, marketing, and data entry.

Table A.1: Top Source Countries, Destination Countries, and Professional Categories

Rank	Destination Country		Source Country		Professional Category	
	Name	Share (%)	Name	Share (%)	Name	Share (%)
1	India	35.35	United States	27.45	Design, Media, and Architecture	36.14
2	Pakistan	25.85	Australia	10.00	Websites, IT, and Software	34.45
3	Bangladesh	12.29	India	9.53	Writing and Content	8.63
4	Egypt	3.52	United Kingdom	8.10	Sales and Marketing	5.42
5	United States	1.68	Canada	5.48	Data Entry and Administration	5.30
	Other	21.31	Other	39.44	Other	10.07

Notes. This table reports the largest source countries, destination countries, and professional categories for cross-border digital service outsourcing observed in the platform data. Shares are computed as percentages of total completed cross-border contracts over the sample period.

Table A.2 reports summary statistics for the skill-level analysis. The unit of observation is a platform-defined skill category. Contract value is the total US dollar value of completed contracts within each skill over the sample period, converted from local currencies using contemporaneous exchange rates. Number of contracts, workers, and buyers count distinct completed transactions, unique individuals completing at least one contract, and unique clients posting at least one contract within each skill, respectively. AI automation exposure is the skill-level index described in Section 3. The distributions are highly right-skewed, with means substantially exceeding medians across all activity measures, reflecting the concentration characteristic of digital labor markets.

C Construction of Generative AI Exposure Measure

This appendix documents the construction of our skill-level AI exposure measure. Measuring automation exposure to generative AI is inherently challenging. Existing measures differ

Table A.2: Summary Statistics: Skill-Level Aggregates

	Mean	SD	P10	P50	P90
Contract value (000s USD)	170.0	1,000.0	0.2	6.1	200.0
Number of contracts	593.0	3,785.2	2.0	21.0	583.0
Number of workers	110.6	358.8	2.0	15.0	253.0
Number of buyers	401.6	2,381.5	2.0	18.0	451.0
AI automation exposure	0.69	0.12	0.56	0.72	0.84

Notes: Summary statistics for 2,277 skill categories with at least one completed contract during January 2022–October 2022. Contract value is the total USD value of completed contracts within each skill, reported in thousands. AI automation exposure is the skill-level index constructed in Section 3, normalized to [0,1].

along at least three dimensions: the unit of analysis (tasks, skills, or occupations), the underlying concept (technological capability versus realized usage), and the mapping between AI capabilities and economic activities. Because our paper studies trade in tasks mediated through an online labor platform where production is organized around narrowly defined skills rather than occupations, the choice of exposure measure is central to identification and interpretation. Our baseline analysis uses a capability-based, skill-level exposure index constructed by eliciting generative AI's ability to perform the tasks associated with each of the platform's 2,700 skill categories.

This appendix describes the elicitation procedure; the stability of exposure scores across repeated elicitations and alternative prompt specifications; comparisons with existing occupation-level indices; and the construction of AI-complementary and interactive intensity measures.

C.1 Elicitation Procedure

Prompt design. For each skill, we construct a prompt that presents the skill name and definition and then requests a quantitative assessment of generative AI's capability to perform the skill.[39] The full prompt template is:[40]

```
System Role: You are a generative AI scholar with deep expertise in digitally
delivered services and repositories of knowledge across skill domains. Your
task is to evaluate individual skills and produce indices that capture generative
AI's ability to automate the skills.

Skill: [SKILL_NAME]
Broad Category: [CATEGORY]
```

[39] Skills requiring any physical presence or offline execution are excluded.

[40] We also considered alternative prompt variations as discussed below, and verified that exposure rankings are robust to prompt wording.

```
Definition: [SKILL_DEFINITION]

On a scale from 0 to 1, how well can generative AI perform a task related
to the skill [SKILL_NAME] compared to a human, with minimal human input?

Please note: You have to perform the actual task with minimal human input.

  • 0 = AI cannot meaningfully perform any of these tasks
  • 0.5 = AI can perform these tasks at roughly half the quality or completeness
    of a human
  • 1 = AI can fully perform these tasks at human-level quality

Response:
```

API parameters. We query the OpenAI API using the following parameters to ensure reproducibility and minimize stochastic variation:

- **Model:** `gpt-3.5-turbo` (baseline); `gpt-4` and `gpt-4o` (robustness)
- **Temperature:** 1 (0: deterministic decoding; 2: most creative)
- **Conversation history:** None (each query is independent)

Stability across runs. To reduce stochastic variations, we repeat each prompt 100 times for each skill and average the responses. The standard error of each skill's mean exposure is approximately 0.008 at the median, an order of magnitude smaller than the cross-skill standard deviation of mean exposure scores (0.12). To put these magnitudes in context: a skill measured at 0.60 exposure has a 95 percent confidence interval of approximately 0.58–0.62. Measurement precision is therefore far finer than the cross-skill variation we exploit for identification. Any remaining measurement error is uncorrelated with true exposure and attenuates estimated effects toward zero.

C.2 Robustness to Alternative Language Models

Our baseline exposure measure is constructed using GPT-3.5, the model available at widespread commercial adoption during our sample period. To assess robustness to alternative language models, we recompute exposure scores using GPT-4 (released March 2023) and GPT-4o (released May 2024).

Table A.3 reports the comparison by reporting pairwise rank correlations across all three model versions. Exposure rankings derived from GPT-3.5 are highly correlated with those

from GPT-4 and GPT-4o (0.63 and 0.56, respectively), and correlations are even stronger between GPT-4 and GPT-4o (0.75).

Table A.3: Rank Correlations of Skill-Level Exposure Scores Across Model Versions

	GPT-3.5	GPT-4	GPT-4o
GPT-3.5	1.00	0.63	0.56
GPT-4		1.00	0.75
GPT-4o			1.00

Notes. Table reports pairwise Spearman rank correlations of skill-level automation exposure scores constructed using different large language model versions. Exposure scores are computed for all platform skill categories and normalized to the unit interval.

Figure A.1 compares skill-level exposure scores generated using GPT-3.5 to those recomputed using GPT-4. Each point corresponds to one of the platform's 2,700 skill categories, with both scores normalized to the unit interval. The figure reveals a strong positive relationship between the two measures, with a Pearson correlation of 0.61 and a Spearman rank correlation of 0.63. This close alignment indicates that the relative ranking of skills by automation exposure is largely stable across model generations.

Importantly, this consistency is not mechanical. GPT-4 and GPT-4o differ from GPT-3.5 along several dimensions, including reasoning depth, instruction following, and generalization; yet exposure rankings remain highly correlated. This suggests that our baseline measure captures the underlying technological feasibility of task automation rather than the idiosyncratic features of a particular model snapshot. Skills assessed as highly automatable by GPT-3.5 tend also to be assessed as highly automatable by more advanced models, even as overall capability expands.

At the same time, the dispersion around the 45-degree line in Figure A.1 is informative. Differences across models are most pronounced among mid-exposure skills, where automation feasibility is inherently ambiguous and more sensitive to incremental improvements in model capabilities. By contrast, skills at the extremes, those that are clearly automatable or clearly resistant, exhibit much tighter alignment across models.

We therefore use GPT-3.5 as our baseline model for two reasons. First, GPT-3.5 corresponds closely to the technological frontier that was widely accessible to firms at the onset of generative AI diffusion. It underpinned the initial release of ChatGPT in November 2022 and reflects the capabilities firms and workers faced when adjustment began. Second, GPT-3.5 provides a conservative benchmark for automation feasibility. To the extent that newer models expand the set of automatable tasks, estimates based on GPT-3.5 are likely to underestimate, rather than overestimate, the longer-run scope of AI-driven reorganization. For

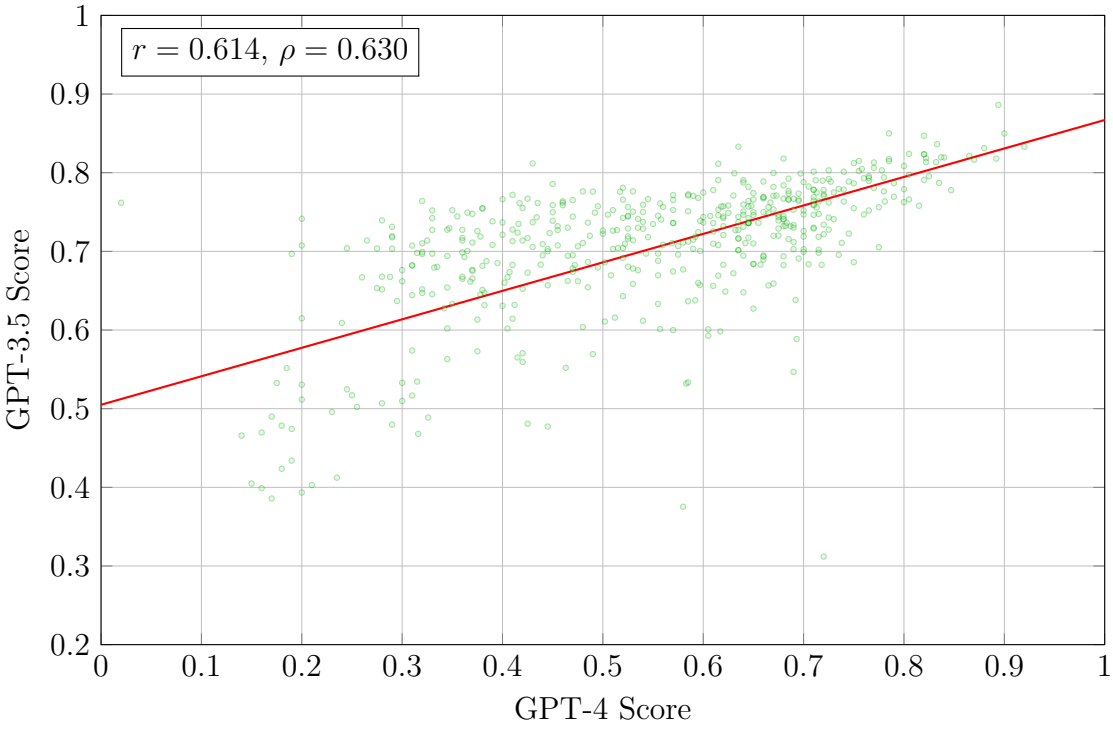


Figure A.1: Comparison of Skill-Level Exposure Scores: GPT-3.5 vs. GPT-4

Notes. Each point corresponds to a platform-defined skill category. The horizontal axis reports exposure scores constructed using GPT-4, and the vertical axis reports exposure scores constructed using GPT-3.5. Scores are normalized to the unit interval.

these reasons, GPT-3.5 is well-suited for studying the early, equilibrium-shaping effects of generative AI on international outsourcing.

C.3 Validation with Internal and External Alternative Indices

Our exposure measure is designed to capture the technological scope for automation facing firms when deciding whether to outsource, automate, or reorganize tasks. A natural concern is whether our findings depend on this particular construction. To address this concern, we compare our measure with both internal alternative and leading external exposure measures developed in prior work.

First, we construct an alternative exposure measure based on AI-assisted task execution rather than full automation. This measure captures settings in which human input remains necessary, but generative AI can substantially augment productivity. The key difference from our baseline measure is that the prompt asks how well generative AI can "assist an individual in performing the tasks" instead of "perform these tasks relative to a human, with minimal human input". The two measures are highly correlated ($r = 0.87$), indicating that skills where AI can substitute for human labor largely overlap with skills where AI can augment it. This correlation suggests that the distinction between automation and augmentation is more a matter of degree than of kind, at least at the current technological frontier. We use the automation-focused measure as our baseline because it more directly captures the margin relevant to outsourcing decisions: whether firms can replace offshore labor with AI, not whether they can make that labor more productive.

In addition, we compare our skill-level measure against two prominent occupation-level indices. First, we use the exposure measures developed by Eloundou et al. (2024), who classify O*NET tasks according to whether large language models can assist with or fully automate them. Their α measure captures the share of an occupation's tasks that AI can assist with; their β measure captures tasks AI can perform with minimal human oversight. Second, we use the Anthropic AI Exposure Index, which measures realized AI usage patterns across occupations based on anonymized conversations with Claude.[41]

Each platform skill is matched to relevant O*NET occupations through a four-layer cascade. First, the skill's category (e.g., "Websites, IT & Software") is used to filter the 923 occupations down to only those in related O*NET broad groups (e.g., Computer and Mathematical + Architecture and Engineering). Next, a curated keyword dictionary checks whether the skill name contains a recognized term (E.g., "python", "accounting") and, if so, selects only occupations whose titles contain the corresponding occupation keywords within

[41] See https://www.anthropic.com/research/economic-index.

that filtered pool. If no keyword rule fires, a token-overlap method tokenises both the skill name and each candidate occupation title (after lowercasing, removing punctuation, and stripping stopwords), and accepts any occupation sharing at least one meaningful token ($length > 3$). If neither the keyword dictionary nor token overlap produces any specific matches, the skill is assigned the average exposure of occupations in its domain-filtered pool.

Table A.4 reports correlations between our measure and the external indices. Our skill-level measure correlates positively with all three benchmarks. The Spearman correlation with Eloundou et al. (2024)'s β (automation potential) is 0.53; with their α (assistance potential), 0.67; and with the Anthropic index, 0.54. Pearson correlations show a similar pattern. These correlations indicate meaningful alignment: skills that our measure identifies as highly exposed also tend to fall within occupations that external indices classify as exposed.

Table A.4: Correlation with External Exposure Indices

External Measure	**Spearman ρ**	**Pearson r**
Eloundou et al. β (automation)	0.53	0.67
Eloundou et al. α (assistance)	0.67	0.54
Anthropic AI Exposure Index	0.54	0.52

Notes: Table reports correlations between our skill-level exposure measure and occupation-level indices from prior work.

At the same time, substantial residual variation remains, which is expected given that the measures are constructed at different levels of aggregation and designed for different purposes. The external indices operate at the occupation level—the appropriate unit for economy-wide workforce analysis—and necessarily average across the heterogeneous skills bundled within each occupation. Our measure, constructed at the platform's skill-category level, is designed to capture within-occupation variation that is relevant for understanding outsourcing decisions but occupation-level indices are not intended to measure. Within "Software Developers," for instance, our measure distinguishes Python scripting (high exposure) from systems architecture (lower exposure), a distinction that matters for predicting which contracts are displaced but that falls below the resolution of occupation-level taxonomies. The documented correlations across these conceptually distinct measures provide stronger validation than would perfect agreement, which could instead suggest redundancy.

Figure A.2 illustrates the complementary granularity of the skill-level approach. Each panel plots the distribution of our skill-level exposure scores (hollow circles) for skills mapped to a given O*NET occupation, alongside the corresponding occupation-level score from Eloundou et al. (2024) (filled circle).

Two patterns emerge. First, within-occupation dispersion is substantial: skills mapped

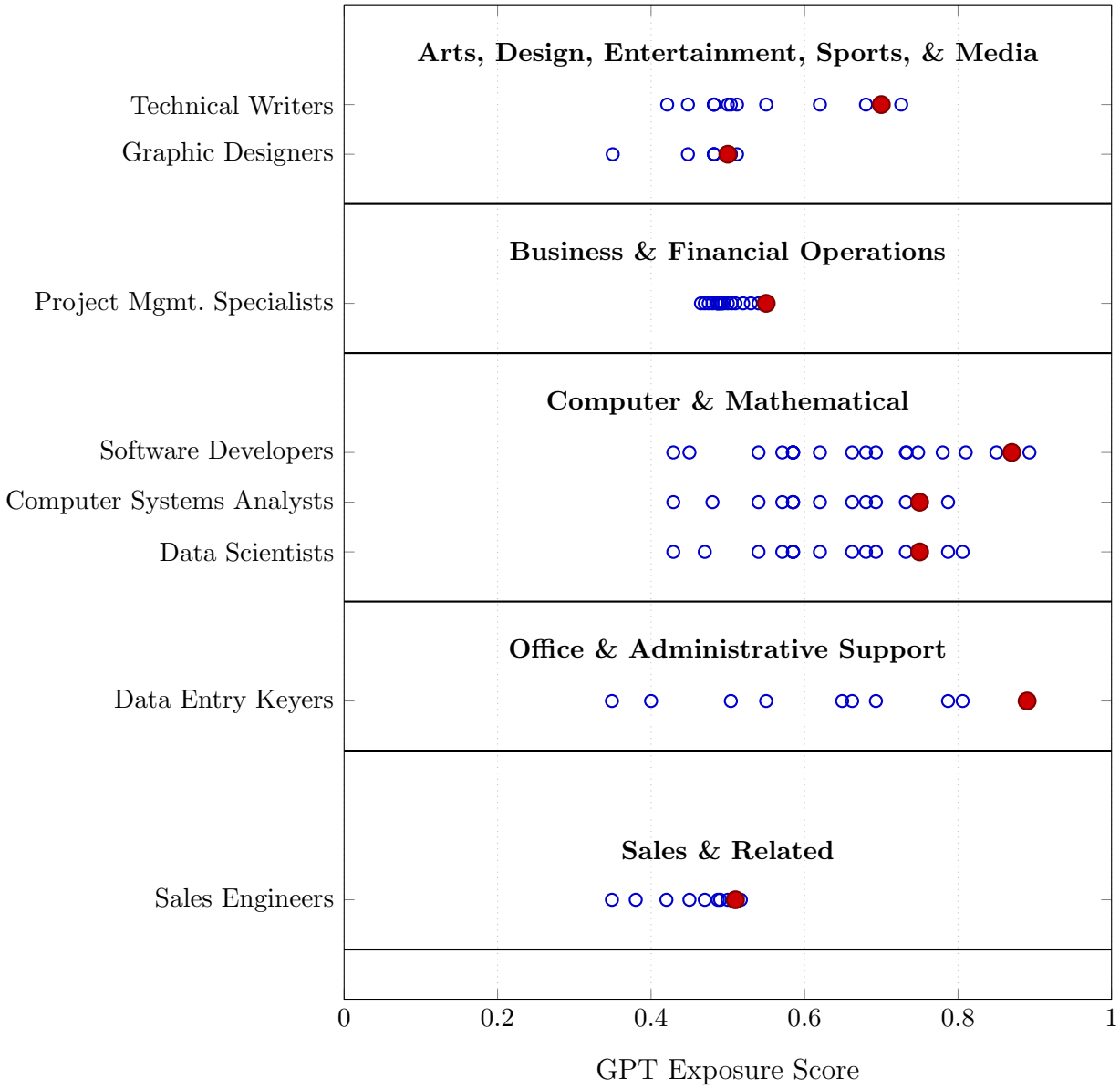


Figure A.2: Within-occupation dispersion of skill-level GPT exposure scores

Notes: This figure displays the distribution of skill-level GPT exposure scores within selected O*NET occupations. Hollow blue circles represent individual skill exposure scores mapped to each occupation. Filled red circles represent occupation-level exposure scores from Eloundou et al. (2024). Occupations are grouped by major O*NET category.

to Software Developers span nearly the full exposure range (0.43 to 0.89), and similar variation appears within Computer Systems Analysts, Data Scientists, and Data Entry Keyers. A single occupation-level score, appropriate for characterizing the occupation as a whole, compresses this variation into a point estimate, treating a COBOL programmer and a Python developer as equally exposed to automation.

Second, for occupations that bundle diverse activities, such as Software Developers, the occupation-level score tends to sit near the upper end of the within-occupation distribution. This is a natural consequence of aggregation: occupation-level assessments reflect the most prominent tasks within a job title, which may not be representative of the full range of constituent skills. By contrast, occupations with narrower task content, such as Project Management Specialists, exhibit tight clustering, and the occupation-level score provides an effective summary.

These patterns clarify why re-estimating our main specifications with occupation-level indices produces qualitatively similar but attenuated coefficients (Appendix D.5): aggregation compresses the within-occupation variation that drives differential outsourcing responses in our data. This is not a limitation of the occupation-level measures, which are designed for broader labor market analysis, but rather an illustration of why skill-level measurement is necessary for our specific empirical setting. At the same time, the qualitative stability of results across all alternative measures suggests that the findings reflect genuine exposure-driven displacement rather than artifacts of a particular measurement approach.

C.4 AI Intensity and Interactive Intensity

AI intensity. We measure AI intensity as the share of contracts within each skill whose project titles or descriptions reference AI-related work—for example, tasks explicitly involving training, deploying, monitoring, or integrating AI tools into production workflows. This measure captures the skill's scope for integrating generative AI as a complement to human labor, rather than its vulnerability to substitution.

Specifically, we classify a project as AI-related if its title or description contains any of the following terms or phrases:

- Explicit AI references: "artificial intelligence," "machine learning," "deep learning," "neural network," "NLP," "computer vision"
- Model operations: "model training," "model deployment," "MLOps," "model monitoring"

- Integration activities: "AI integration," "chatbot development," "prompt engineering," "LLM," "GPT"
- Data preparation: "data labeling," "annotation," "training data"

For each skill s, AI intensity is defined as:

$$\text{AIIntensity}_s = \frac{\#\text{ projects in skill } s \text{ with AI-related keywords}}{\#\text{ projects in skill } s}, \tag{A-22}$$

computed over all projects posted in the pre-ChatGPT period (January–October 2022).

Interactive intensity. We measure the extent to which tasks within each skill require interpersonal interaction, client communication, or soft skills that are less amenable to automation. For each skill, we compute the share of associated projects (in the pre-period) whose titles or descriptions contain terms indicating interactive or relational activities.

The keyword dictionary includes:

- Communication: "meeting," "call," "discuss," "collaborate," "coordinate"
- Client interaction: "client," "stakeholder," "feedback," "revision," "consultation"
- Creativity and judgment: "creative," "brainstorm," "strategy," "advise," "recommend"
- Presentation: "present," "pitch," "demo," "workshop"

For each skill s, interactive intensity is defined as:

$$\text{Interactive}_s = \frac{\#\text{ projects in skill } s \text{ with interactive keywords}}{\#\text{ projects in skill } s} \tag{A-23}$$

computed over all projects posted in the pre-ChatGPT period (January–October 2022).

Table A.5 reports summary statistics for the three skill-level measures used in the analysis: exposure to generative AI automation, interactive intensity, and AI-related skill content. The AI exposure measure exhibits substantial dispersion across skills, with a standard deviation of 0.12 and a wide interdecile range, indicating meaningful heterogeneity in the technological feasibility of automation even within narrowly defined digital tasks. By contrast, interactive intensity is lower on average but highly skewed, reflecting that a subset of skills involves frequent client communication and coordination, while many tasks remain largely transactional. AI-specific skills remain uncommon, accounting for 5 percent of skills.

Although these measures are related conceptually, their correlations are modest. AI exposure is only weakly positively correlated with AI-related skill status, reflecting the fact that skills associated with AI span a wide range of automation potential, from routine data preparation to specialized model development and oversight. AI exposure is also positively and weakly correlated with interactive intensity. The magnitudes of these correlations are small (0.05-0.15), confirming that the three measures capture distinct dimensions of skill content rather than redundant constructions. This variation is critical for identification: it allows us to distinguish between skills that are automatable, skills that are specific to AI, and skills that are more dependent on human interaction, and to trace how generative AI reshapes outsourcing, buyer behavior, and worker adjustment along each of these margins.

Table A.5: Summary Statistics: AI Exposure Measure

Variable	**Mean**	**SD**	**P10**	**Median**	**P90**	**N**
AI Automation Exposure	0.69	0.12	0.54	0.70	0.83	2,458
AI Intensity	0.01	0.06	0	0	0.05	2,388
Interactive Intensity	0.23	0.25	0	0.21	0.52	2,260

Notes: Table reports summary statistics for skill-level measures. Skills with at least one completed contract in the sample period are included. AI Automation Exposure is the exposure measure capturing GPT's assessed ability to substitute for workers in that skill, normalized to the unit interval. Interactive Intensity is the share of pre-period projects containing interactive keywords. AI Intensity is the share of pre-period projects within each skill whose titles or descriptions reference AI-related terms.

D Sensitivity Analysis

D.1 Functional Form Sensitivity: Binned Exposure

This sensitivity analysis assesses whether the aggregate contraction in outsourcing documented in the main text depends on functional-form assumptions embedded in the continuous exposure specification. The baseline difference-in-differences model assumes that post-GPT outcomes vary linearly with exposure to generative AI. Economic theory, however, does not require such linearity: automation may operate through threshold effects, whereby tasks become abruptly unprofitable to outsource once AI capabilities exceed a certain level.

To relax this assumption, we discretize the continuous exposure score into quartiles based on its pre-GPT distribution and re-estimate the aggregate specification. The lowest-exposure quartile serves as the omitted category. This approach allows for non-linear responses and guards against the possibility that baseline results are driven by a small number of highly

exposed skills. The regressions are estimated at the skill–week level using the same outcomes, fixed effects, and clustering structure as in the main analysis.

Table A.6: Aggregate Effects of Generative AI on Outsourcing to Developing Countries: Binned Exposure

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Buyers)	(4) Log(Workers)
PostGPT × Bin 2	-0.194***	-0.065**	-0.071***	-0.069***
	(0.052)	(0.028)	(0.027)	(0.025)
PostGPT × Bin 3	-0.171***	-0.090***	-0.096***	-0.098***
	(0.047)	(0.026)	(0.025)	(0.023)
PostGPT × Bin 4	-0.255***	-0.144***	-0.146***	-0.147***
	(0.045)	(0.025)	(0.024)	(0.022)
Skill FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
R^2	0.637	0.886	0.891	0.884
Observations	88,800	90,045	90,045	90,045

Notes. The dependent variables are measured at the skill-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a skill-level exposure measure capturing GPT's assessed ability to substitute for workers in that skill, normalized to the unit interval. Exposure scores are divided into quartiles based on the pre-GPT distribution; Bin 1 (lowest exposure) is the omitted category. All specifications include skill fixed effects and week fixed effects. Standard errors are clustered at the skill level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results in Table A.6 show a clear and monotonic gradient in the post-GPT contraction of outsourcing across exposure bins. Relative to the lowest-exposure skills, higher-exposure bins experience progressively larger declines in outsourcing value, contract volume, buyer participation, and worker participation. Even moderately exposed skills (Bin 2) exhibit statistically significant declines across all outcomes, indicating that the effects of generative AI are not confined to the extreme upper tail of exposure. This exercise offers additional support to the robustness of the baseline findings documented in 5 of the main paper.

D.2 Inverse Hyperbolic Sine Transformation

The main specification in Section 5 uses log-transformed dependent variables, which drops skill-week observations with zero values. This approach raises a potential concern: if AI exposure causes some skill-weeks to have zero activity, then dropping zeros may understate the true displacement effect by excluding the most severely affected observations.

Table A.7 addresses this concern using the inverse hyperbolic sine (IHS) transformation. The IHS transformation approximates the natural logarithm for large values of y but is de-

Table A.7: Aggregate Effects of Outsourcing to Developing Countries: IHS Transformation

	(1) IHS(Value)	(2) IHS(Contracts)	(3) IHS(Buyers)	(4) IHS(Workers)
Post-GPT × Exposure	−1.236***	−0.523***	−0.512***	−0.496***
	(0.141)	(0.045)	(0.044)	(0.040)
Skill FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
R^2	0.685	0.869	0.871	0.865
Observations	328,536	328,536	328,536	328,536

Notes. The dependent variables are measured at the skill-week level and transformed using the inverse hyperbolic sine (IHS). *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a skill-level exposure measure capturing GPT's assessed ability to substitute for workers in that skill, normalized to the unit interval. All specifications include skill fixed effects and week fixed effects. Standard errors, clustered at the skill level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

fined at zero, allowing us to retain all observations including skill-weeks with no completed contracts. The results are quantitatively similar to the baseline log specification. The coefficients on contract volume, active buyers, and active workers are comparable in magnitude and remain highly significant. The consistency across specifications indicates that the baseline findings are not driven by selective attrition of zero-value observations. Specifically, the similar effects offer two interpretations. First, complete market exit at the skill-week level is relatively rare; most of the displacement operates on the intensive margin (reduced activity within active skill-weeks) rather than the extensive margin (skills dropping to zero). Second, to the extent that zeros do occur, they are distributed similarly across high- and low-exposure skills conditional on fixed effects, so their exclusion does not systematically bias the baseline estimates.

D.3 Controlling for Broad Skill Category Shocks

Our baseline specification exploits variation in AI exposure across all skill categories. However, a potential concern is that skills with high AI exposure are concentrated in particular broad categories (e.g., "Websites, IT & Software") that may follow differential trends for reasons unrelated to automation, such as sector-specific demand cycles or shifts in client preferences. If these category-level trends correlate with our exposure index, the baseline estimates could capture sectoral dynamics rather than skill-level responses to generative AI.

To address this, we augment Equation (1) with broad category-by-month fixed effects,

which absorb all time-varying shocks common to skills within the same category. Identification then relies exclusively on within-category variation: comparing more and less exposed skills within the same field over time.

Table A.8: Aggregate Effects on Outsourcing to Developing Countries: Controlling for Broad Category Shocks

	Log(Value)	Log(Contracts)	Log(Buyers)	Log(Workers)
Post-GPT × Exposure	-0.739^{***}	-0.554^{***}	-0.549^{***}	-0.568^{***}
	(0.143)	(0.083)	(0.080)	(0.075)
Skill FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Category × Month FE	Yes	Yes	Yes	Yes
Observations	88,795	90,039	90,039	90,039
R^2	0.641	0.889	0.894	0.887

Notes: This table reports estimates of Equation (1) with the addition of broad skill category-by-month fixed effects. The dependent variables are measured at the skill-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. *Exposure* is a skill-level exposure measure capturing GPT's assessed ability to substitute for workers in that skill, normalized to the unit interval. Standard errors clustered at the skill level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8 reports the results. The estimates are quantitatively similar to those reported in Table 2: the slightly attenuated magnitudes indicate that a small portion of the baseline effect reflects differential trends across broad categories. The remaining estimates are large and highly significant across all four outcomes, confirming that the contraction in outsourcing is driven by fine-grained, skill-level variation in automation exposure rather than by sectoral composition.

D.4 Major Model Releases

A natural question is whether the displacement we document is driven specifically by OpenAI's products or by the broader expansion of the generative AI frontier. To investigate, we incorporate into the analysis major model releases across providers, including Anthropic's Claude family and Google's Gemini, and interact each with skill-level exposure. As shown in Figure A.3, significant incremental displacement effects emerge following releases by multiple providers (GPT-3.5, Claude 3, GPT-4o), not exclusively OpenAI models. This pattern is consistent with buyers responding to the aggregate capability frontier rather than to any single provider: as competing models reach and extend the automation threshold, additional outsourced tasks become candidates for substitution.

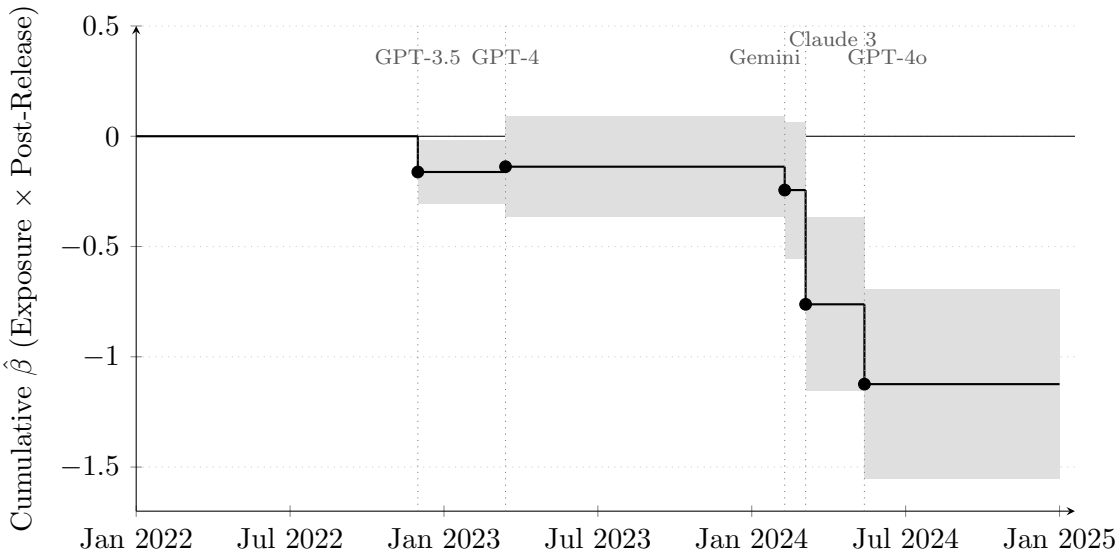


Figure A.3: Cumulative Effect of AI Model Releases on Contract Volume

Notes: This figure plots the cumulative differential effect of AI exposure on log contract count, based on estimates from a multi-release generalization of Equation (1). The specification replaces the single Post-GPT indicator with separate indicators for weeks following each major model release: GPT-3.5 (November 30, 2022), GPT-4 (March 14, 2023), Gemini (February 8, 2024), Claude 3 (March 4, 2024), and GPT-4o (May 13, 2024). Each indicator is interacted with the skill-level automation exposure score. The cumulative coefficient at each release date equals the sum of all incremental interaction coefficients up to and including that release, capturing the total differential effect of exposure on contract volume at that point in time. Shaded bands represent 95 percent confidence intervals for the cumulative coefficients. All specifications include skill fixed effects and week fixed effects. Standard errors are clustered at the skill level.

D.5 Benchmarking with Alternative Exposure Indices

This appendix section evaluates the robustness of our baseline results to alternative exposure measures discussed in Appendix C.3. The purpose of this exercise is twofold. First, we assess whether the aggregate contraction in cross-border outsourcing documented in the paper is sensitive to the measure of exposure to generative AI. Second, we clarify how different exposure measures correspond to distinct economic concepts (capability, task susceptibility, and occupational composition) and why the skill exposure measure is necessary for studying the reorganization of international outsourcing and skill dynamics in our setting.

For each alternative exposure measure, we re-estimate the aggregate difference-in-differences specification, Equation (1), at the skill–week level. Outcomes include the log total value of completed contracts, the log number of contracts, the log number of active buyers, and the log number of active workers in a skill subcategory. All specifications include skill fixed effects and week fixed effects, and standard errors are clustered at the skill level.

Table A.9 reports the results. Across all exposure measures, we find consistent evidence that generative AI is associated with a contraction in cross-border outsourcing to developing countries. For each measure, higher exposure predicts statistically significant declines in outsourcing value, contract volume, buyer participation, and worker participation following the introduction of generative AI. The sign and relative ordering of effects on the outcomes are remarkably stable across measures, indicating that our central qualitative conclusion—that generative AI reduces outsourcing in more automatable tasks—is not an artifact of our exposure construction.

At the same time, the magnitude of estimated effects differs across exposure measures in economically meaningful ways. The largest effects, evaluated at the mean exposure levels, are obtained using our baseline capability-based exposure score. Measures from Eloundou et al. (2024) and Anthropic yield smaller effects, though they remain statistically significant. This attenuation is expected given the coarser level of matching and aggregation: mapping occupational exposure scores or scores based on a different task taxonomy to platform skills necessarily introduces measurement error and averages cross-skill heterogeneity.

Importantly, differences in magnitude should not be interpreted as contradictions across measures. Instead, they reflect differences in what each measure captures. Occupation-based indices are constructed to assess exposure of jobs or workers in conventional labor markets and therefore average over heterogeneous tasks within occupations. This aggregation also permits adjustment across tasks within the same occupation, which mechanically attenuates localized, within-skill effects of generative AI. Our setting, in contrast, is a marketplace for narrowly defined digital tasks, where reallocation occurs primarily within occupations and across fine-grained skills. While the sensitivity analysis shows that the aggregate contraction

Table A.9: Benchmarking with Alternative Exposure Indices

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Buyers)	(4) Log(Workers)
Panel A. Assistance Exposure Score				
Post-GPT × Exposure	-0.686***	-0.473***	-0.472***	-0.456***
	(0.109)	(0.065)	(0.063)	(0.060)
Observations	88,800	90,045	90,045	90,045
R^2	0.637	0.887	0.892	0.884
Panel B. Alpha Exposure (Eloundou et al., 2024)				
Post-GPT × Exposure	-0.326**	-0.226***	-0.242***	-0.208***
	(0.082)	(0.053)	(0.052)	(0.047)
Observations	88,692	89,937	89,937	89,937
R^2	0.635	0.886	0.891	0.883
Panel C. Beta Exposure (Eloundou et al., 2024)				
Post-GPT × Exposure	-0.291**	-0.187**	-0.212***	-0.173**
	(0.135)	(0.083)	(0.081)	(0.074)
Observations	88,692	89,937	89,937	89,937
R^2	0.635	0.885	0.891	0.883
Panel D. Anthropic AI Exposure Index				
Post-GPT × Exposure	-1.013***	-0.424***	-0.421***	-0.496***
	(0.208)	(0.128)	(0.126)	(0.117)
Observations	89,098	90,352	90,352	90,352
R^2	0.634	0.885	0.890	0.882
Skill FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes. The dependent variables are measured at the skill-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. Panel A uses the paper's assistance (capability-based) exposure score, constructed by eliciting GPT-3.5's ability to perform tasks associated with each skill and normalized to lie between 0 and 1. Panels B and C use the Alpha and Beta exposure measures from Eloundou et al. (2024), mapped from tasks to a subset of platform skill categories. Panel D uses the Anthropic AI Exposure Index, mapped from broader occupations to platform skill subcategories. All specifications include skill subcategory and week fixed effects. Standard errors are clustered at the skill level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in outsourcing is robust to alternative exposure measures, these external indices cannot replace our baseline measure for our granular buyer and worker analyses. Our measure, by operating at the level of platform-defined skills which directly correspond to observed contracts, payments, and task bundles, is needed to trace reallocation and reskilling across the exact units of production traded in the market.

D.6 Buyer Heterogeneity in Task Reorganization

This subsection examines heterogeneity in buyer-side adjustment to generative AI. Specifically, the specification allows the post-GPT exposure effect to vary with two predetermined buyer characteristics measured before November 30, 2022: pre-GPT outsourcing expenditure, which proxies for the scale of a buyer's activities on the platform, and pre-GPT skill breadth, which captures the diversity of tasks previously outsourced. Because both variables are fixed in the pre-treatment period and all regressions include buyer fixed effects, identification comes from within-buyer changes over time, interacted with differential exposure to generative AI.

Table A.10: Shifts in Outsourcing Demand: Heterogeneity by Buyer Scale and Scope

	(1) Log(Value)	(2) Log(Contracts)	(3) Log(Mean Value)
PostGPT × Exposure	1.466***	-0.022	1.484***
	(0.246)	(0.045)	(0.241)
PostGPT × Exposure × log(PreGPT Expenditure)	-0.404***	-0.005*	-0.397***
	(0.012)	(0.003)	(0.012)
PostGPT × Exposure × log(PreGPT Skills)	0.541***	-0.040***	0.578***
	(0.028)	(0.008)	(0.027)
Buyer FE	Yes	Yes	Yes
Source Country × week FE	Yes	Yes	Yes
R^2	0.596	0.356	0.608
Observations	92,583	94,411	92,583

Notes. The dependent variables are measured at the buyer-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. log(PreGPT Expenditure) and log(PreGPT Skills) are measured over the employer's pre-GPT outsourcing history and proxy for pre-period outsourcing scale and scope. All specifications include employer fixed effects, week fixed effects, and source-country × week fixed effects. Standard errors are clustered at the employer level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Consistent with Section 6, the results in Table A.10 again reveal a pattern of outsourcing reallocation. For exposed buyers at baseline levels of pre-GPT scale and scope, total outsourcing value and mean contract value increase significantly after ChatGPT's introduction, while the number of completed contracts does not. This pattern indicates that exposed

buyers who remain active shift toward higher-value engagements rather than expanding outsourcing along the extensive margin, consistent with a withdrawal from routine tasks and a reorientation toward work where human input remains valuable.

Heterogeneity by pre-GPT expenditure indicates that this value upgrading is attenuated among larger buyers. The negative interaction between exposure and pre-period expenditure implies that buyers with greater pre-GPT spending experience larger declines in outsourcing contract volume and smaller increases in outsourcing value and mean contract value. One interpretation is that larger buyers have greater internal capacity to absorb exposed tasks using AI, reducing their reliance on external contractors when automation becomes feasible.

In contrast, pre-GPT skill breadth amplifies value upgrading among exposed buyers. Buyers that previously sourced a wider range of skills exhibit larger post-GPT increases in both total value and mean contract value, even as they source fewer contracts. This pattern suggests that diversity in skill requirement portfolios facilitates consolidation into fewer, higher-value engagements when generative AI reduces the profitability of routine outsourcing.

The above evidence shows that generative AI induces selective reorganization rather than uniform contraction on the demand side. Exposure to AI shifts outsourcing toward higher-value work among buyers with the organizational capacity to adapt, while buyers with greater scale are more likely to substitute away from outsourced production altogether.

D.7 Heterogeneity across Major Outsourcing Destinations

Table A.11 examines whether the aggregate contraction in outsourcing documented in the main analysis is broadly shared across major developing-country destinations or driven by a narrow subset of countries. We re-estimate the baseline skill–week difference-in-differences specification separately for the five largest destination countries on the platform—India, Pakistan, Bangladesh, Egypt, and Sri Lanka using identical fixed effects, exposure measures, and timing.

Two patterns stand out. First, the negative relationship between generative-AI exposure and outsourcing activity is present across nearly all major destinations. For India, Pakistan, Bangladesh, and Sri Lanka, higher exposure is associated with statistically significant declines in total contract value, contract volume, and active workers following the introduction of generative AI. This consistency indicates that the aggregate results are not driven by a single country or idiosyncratic institutional context, but instead reflect a common technological shock affecting digitally tradable tasks across multiple labor markets.

Second, the magnitude of the response varies systematically across countries. The contraction is largest in Bangladesh and India, countries that, prior to generative AI, specialized

Table A.11: Aggregate Effects across Major Destinations

	(1) India	(2) Pakistan	(3) Bangladesh	(4) Egypt	(5) Sri Lanka
Panel A. Log(Value)					
PostGPT × Exposure	-0.657***	-0.645***	-1.112***	-0.016	-0.600*
	(0.182)	(0.210)	(0.335)	(0.277)	(0.359)
Observations	47,905	49,218	28,985	18,252	6,818
R-squared	0.585	0.555	0.506	0.382	0.452
Panel B. Log(Contracts)					
PostGPT × Exposure	-0.390***	-0.360***	-0.615***	-0.212*	-0.692***
	(0.093)	(0.130)	(0.182)	(0.123)	(0.195)
Observations	48,911	50,252	29,500	18,721	7,033
R-squared	0.869	0.853	0.810	0.667	0.723
Panel C. Log(Workers)					
PostGPT × Exposure	-0.437***	-0.364***	-0.673***	-0.240*	-0.391***
	(0.084)	(0.099)	(0.142)	(0.135)	(0.140)
Observations	48,911	50,252	29,500	18,721	7,033
R-squared	0.863	0.854	0.816	0.670	0.687
Skill FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variables are measured at the destination-country-skill-week level. *Value* is the total dollar value of completed contracts, converted from local currency using contemporaneous exchange rates. *Contracts* is the total number of completed contracts. *Buyers* is the number of distinct clients posting at least one contract. *Workers* is the number of distinct individuals completing at least one contract. *Post-GPT* is an indicator equal to one for weeks after the introduction of ChatGPT on November 30, 2022. All regressions include skill and week fixed effects. Standard errors are clustered at the skill level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

most heavily in routine, codifiable digital services where wage arbitrage played a central role. In contrast, effects are smaller and less precisely estimated for Egypt, where baseline outsourcing volumes are lower and skill composition is more heterogeneous. This pattern is consistent with the task-based interpretation developed in the paper: generative AI disproportionately erodes demand for tasks that formed the core of comparative advantage in certain developing-country labor markets, rather than uniformly affecting all destinations.

Importantly, the similarity of responses across value, contracts, and workers within each country suggests that these effects reflect genuine contraction in market activity rather than substitution across participants within countries. The results reinforce the interpretation that generative AI operates as a geography-unbounded form of automation, compressing cross-border arbitrage in routine digital tasks, and that its impact is felt most strongly in countries whose export profiles were most exposed to those tasks.

D.8 Skill Transition Dynamics

This appendix reports a gravity-style analysis of worker skill transitions motivated by the model in Appendix A.3. Section 7.2 documents that workers in high-exposure skills reallocate toward AI-complementary and interaction-intensive activities. Here, we characterize the macro-level structure of these transitions by examining whether aggregate reallocation patterns are shaped by adoption frictions, proxied by skill-pair relatedness, and by destination-skill characteristics that the model predicts should attract transitioning workers.

We construct a transition matrix in which each element ω_{ij} denotes the number of workers who completed at least one contract in skill i prior to November 2022 and at least one contract in skill j during the post-period (December 2022–December 2024). A worker may appear in multiple origin-destination pairs if they supplied multiple skills pre-period or transitioned to multiple destinations post-period.

The gravity equation derived in Appendix A.3.4 predicts that transition flows between skills depend on bilateral switching costs (lower between technologically proximate skills) and on the expected payoff at the destination (higher for skills where human inputs remain complementary to AI). We take this prediction to the data by estimating:

$$\log \omega_{ji} = \alpha_j + \beta_1 \log \text{Relatedness}_{ji} + \beta_2 \log \bar{V}_i + \mathbf{X}_i' \boldsymbol{\gamma} + \varepsilon_{ji}, \tag{A-24}$$

where α_j are origin-skill fixed effects absorbing the mass of workers in each origin skill, their average propensity to transition, and their outside options. The origin fixed effects ensure that destination-skill coefficients identify pull factors, i.e., characteristics of skill i that attract transitioning workers, conditional on leaving any given origin.

The two structural covariates correspond directly to the model's predictions:

- *Skill-pair relatedness* (Relatedness_{ji}): the log number of workers who completed contracts in both skills j and i prior to November 2022, proxying for bilateral switching costs. The model predicts $\beta_1 > 0$.
- *Pre-GPT skill payoff* ($\bar{V}_i$): the log average contract value in destination skill i before November 2022, capturing the expected return to entering that market. The model predicts $\beta_2 > 0$.

The vector $\mathbf{X}_i$ includes two additional destination characteristics that are not primitives of the model but help characterize the direction of reallocation: *interaction intensity* (the share of pre-period contracts in skill i involving communication, coordination, creativity, or client-facing judgment); and *AI intensity* (the fraction of contracts referencing AI-related tasks before November 2022). Standard errors are clustered at the origin-skill level.

Table A.12 reports estimates. The two structural predictions receive strong support. Skill-pair relatedness is the dominant predictor of transitions, with a stable elasticity of 0.75-0.83 across specifications: workers move primarily along technologically proximate pathways. Conditional on relatedness, workers transition toward higher-payoff destinations; the coefficient on pre-GPT skill payoff is positive and highly significant. Controlling for destination payoff reduces the relatedness elasticity from 0.83 to 0.75, indicating that part of the raw relatedness effect operates through correlated payoff levels across proximate skills.

Among the other destination characteristics, the coefficient of interaction intensity is positive and significant. Conditional on proximity and payoff, workers disproportionately reallocate to tasks involving communication, coordination, and judgment, consistent with the broader prediction that human inputs shift toward AI-complementary activities. This result confirms at the skill-pair level the worker-level patterns documented in Section 7.2.

The coefficient of destination-skill AI intensity is positive but insignificant. This asymmetry is interpretable through the model's switching cost mechanism: interaction-intensive tasks draw on broadly transferable competencies accessible to displaced workers along proximate margins, whereas AI-specific tasks demand technical knowledge (e.g., model architectures, API integration) that entails high switching costs for many contractors regardless of the destination payoff. The coexistence of individual-level AI earnings premia (Section 7.3) and limited macro-level transition flows into AI destinations illustrates a general implication of the gravity structure—switching costs can sustain earnings differentials in equilibrium—and suggests that targeted training interventions may be necessary to facilitate adjustment beyond technologically adjacent domains.

Table A.12: Gravity Estimates of Skill Transitions

	Log(Worker Flow)				
	(1)	(2)	(3)	(4)	(5)
Skill-Pair Relatedness	0.828***	0.748***	0.829***	0.828***	0.748***
	(0.008)	(0.006)	(0.007)	(0.008)	(0.006)
Destination Skill —					
Pre-GPT Payoff		0.085***			0.088***
		(0.003)			(0.003)
Interaction Intensity			0.157***		0.330***
			(0.058)		(0.057)
AI Intensity				0.224	0.164
				(0.205)	(0.173)
Origin Skill FE	Yes	Yes	Yes	Yes	Yes
Observations	357,852	357,680	357,852	357,852	357,680
R^2	0.739	0.756	0.740	0.739	0.757

Notes. The dependent variable is the log number of workers transitioning from origin skill i to destination skill j, where a transition is defined as completing at least one contract in skill i before November 2022 and at least one contract in skill j afterward. *Skill-pair relatedness* is the log number of workers active in both skills i and j during the pre-period, capturing supply-side proximity between skill categories. *Pre-GPT skill payoff* is the log average dollar value of contracts in skill j before November 2022, controlling for destination market size. *Interaction intensity* is the share of pre-period contracts in destination skill j involving communication, coordination, creativity, or client-facing judgment (Appendix C.4). *AI intensity* is the fraction of contracts in destination skill j referencing AI-related tasks before November 2022. All specifications include origin skill fixed effects. Standard errors clustered at the origin skill level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$