# Splitting the Stack: Generative AI and the Cross-Country Division of Technical Work

Hitanshu Pandit[1]

October 2026

Click here for latest version

## Abstract

When a firm can have a task done by its workers in several countries, how does it decide where each task is done, and what does that decision mean for workers? I read the job descriptions that the same firms post in the United States, India, the Philippines, Mexico and the United Kingdom and classify each technical job by the task it describes. Within the same firm and year, firms place the building of models and data systems less readily in lower-wage countries than the application of technology. Countries do not climb a single ladder of tasks: India specializes in technical application, the Philippines in functional application and the United Kingdom in product and testing. After ChatGPT's release, the firms' U.S. postings for software developers who code applications declined while those for developers who build models and data systems did not, and I detect no shift in where firms place tasks.

Keywords: trade in tasks, task location, task measurement, generative AI, job postings.
JEL No. J23, J24, F16, F23, O33

[1]Bates College and Northeastern University. Email: hitanshupanditphd@gmail.com. I thank Martin Andersen and participants at the 4th AI and White-Collar Work Conference (Sweden) for helpful comments. All errors are my own.

# 1 Introduction

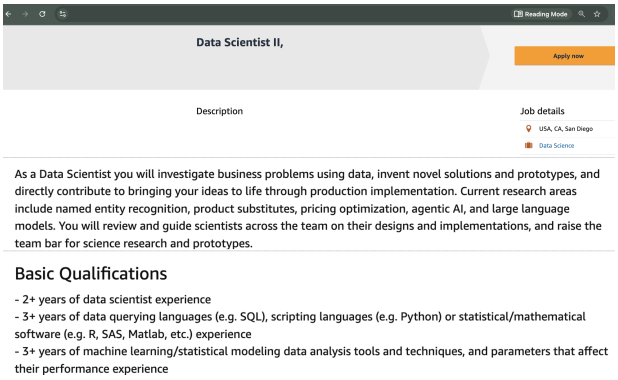
When a firm can have a task done by its workers in any of several countries, how does it decide where each task is done, and what does that decision mean for the workers in each place? The task approach to labor markets (Autor et al. 2003; Acemoglu and Autor 2011) and the theory of trade in tasks (Grossman and Rossi-Hansberg 2008) both answer by asking how work is divided across tasks, and the answer decides which skills each country's workers use, which jobs are open to new entrants and how a new technology reaches the work. Office work has been splitting across countries since communication became cheap (Blinder 2006; Baldwin 2016), and a large part of the U.S. workforce now does work that could, in principle, be delivered over a wire (Blinder 2009; Blinder and Krueger 2013).

Technical work is where these questions now meet. Large firms run their software, data and testing work through offices in many countries, and in the same year one bank or technology company may advertise such jobs in the United States, India, the Philippines, Mexico and the United Kingdom. Technical work is also exposed to generative AI. Language models can perform a large share of the tasks in many white-collar occupations (Eloundou et al. 2024) and raise the productivity of the workers who use them (Brynjolfsson et al. 2025b), and their spread has coincided with weaker demand for early-career workers in exposed U.S. occupations (Westby et al. 2026; Brynjolfsson et al. 2025a). A new technology displaces workers from the tasks it automates while new tasks reinstate labor elsewhere (Acemoglu and Restrepo 2019, 2022); when a firm can place a task in several countries, the technology can also change where the task is done. This paper asks how firms decide where each technical task is done, and whether generative AI changed which tasks firms hire for, where they place them, or both.

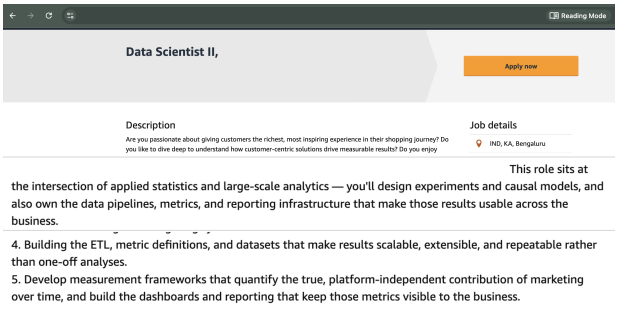
The answer has been hard to see because the unit that moves is hidden in the usual data. Evidence on where white-collar work is done is almost all at the level of the occupation or the industry, and an occupation code is too coarse for technical work. The single code Software Developers covers 39 percent of the firms' Indian technical postings and holds jobs as different as training a recommendation model and maintaining a payroll application. Even a job title can hide the difference. Two "Data Scientist II" postings of one large technology firm, one in San Diego and one in Bengaluru, share a title and a code (Figure 1). The first researches large language models; the second builds the pipelines and dashboards that report experiments. Such pairs are common. When a firm advertises the same title in the same year in both countries, the two countries' postings have a different main task in 26 percent of cases, more than twice the 12 percent disagreement between two random halves of U.S. postings of the same size.

Figure 1: Same firm, same title, different tasks: data scientists of two large technology firms in the United States and India

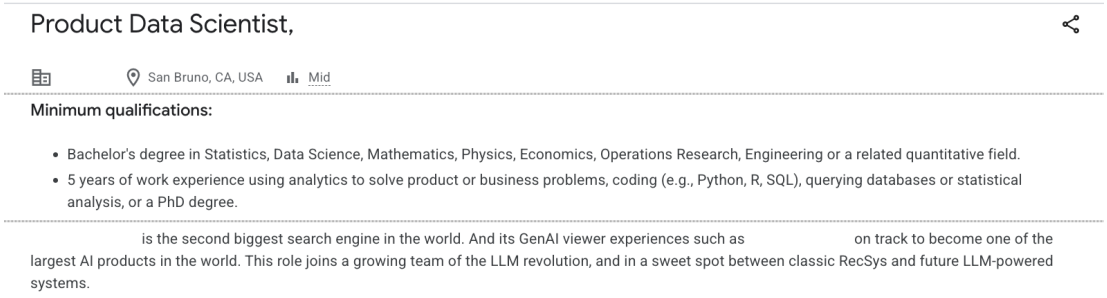
(a) Large technology firm, "Data Scientist II", San Diego



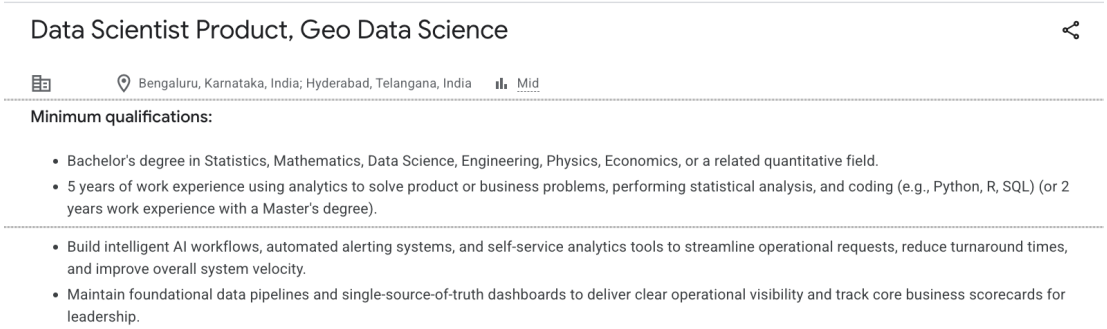

(b) Large technology firm, "Data Scientist II", Bengaluru




(c) Second technology firm, "Product Data Scientist" (Mid), San Bruno




(d) Second technology firm, "Data Scientist Product" (Mid), Bengaluru and Hyderabad




*Notes:* Screenshots of the employers' career sites, saved October 28, 2026. Panels (a) and (b): one large technology firm, same job title; the U.S. posting asks for two or more years of data-scientist experience and the Indian posting for four or more (in its Basic Qualifications, below the excerpt). Panels (c) and (d): a second technology firm; both postings ask for five years of experience and a quantitative degree but describe different tasks. Firm, team and product names are covered

I therefore read the job responsibilities/duties section of each posting and place the job in one of seven parts of the technology stack. Model building and data and machine-learning infrastructure make up *Building*. Technical application, product and testing, functional application (enterprise systems and analytics) and governance and security make up *Application*. The seventh part, general tech, covers the networks, servers and help desks that every firm needs, and serves as the benchmark. The category descriptions contain no AI vocabulary, so a job does not change category because its employer starts mentioning AI. On postings whose titles leave no doubt about the work, the labels agree with the title 90.5 percent of the time.

To read these jobs I extend the model of Grossman and Rossi-Hansberg (2008). Each destination has its own costs of distance, tasks differ in how much graduate labor they need, and firms place individual jobs rather than whole tasks. The result is a gravity equation for jobs inside the firm, and the question it asks is simple: does a firm advertise its Building jobs in a given country as readily as its Application jobs of the same year? Because every comparison holds the firm and the year fixed, the firm's wages, the size of its office in each country and how completely LinkUp covers its career site all cancel out. The estimates describe how firms divide their work; they do not say what caused the division.

The main finding is that firms place tasks selectively, and in the same direction in every lower-wage country. Take two jobs that one firm advertises in the same year, one building models or data systems and one applying them. The Building job's odds of being advertised in India rather than the United States are 18 percent lower than the Application job's, 26 percent lower in Mexico and 37 percent lower in the Philippines. In the United Kingdom the gap is close to zero on the observed labels, though it widens under some corrections for misclassification. The gap is widest for work that most often requires a doctorate. Titles for which more than 30 percent of U.S. postings ask for a PhD, such as applied scientists and machine-learning engineers, have 30 percent lower odds of being advertised in India than titles that never ask for one, and 80 percent lower odds in the Philippines. Firms, in other words, place the tasks that use the scarcest skills where those skills are.

The pattern is richer than the simplest theory allows. If tasks differed only in how hard they are to do at a distance, every country would take them in the same order, like the rungs of one ladder. They do not. Read from the job descriptions, India specializes in technical application, the Philippines in functional application and the United Kingdom in product and testing, and a formal test rejects a single ranking decisively. Doctoral skill accounts for part of the pattern. The Philippines and Mexico, where doctorates are scarce, fit the ordering by doctoral supply, but India does not separate from the United Kingdom in the linear specification, and its penalty is concentrated among the most PhD-intensive titles.

India's record shows what climbing the stack looks like. Between 2019 and 2025, inside the same firms, India's odds of receiving Building work rose relative to general tech, and its odds of receiving Application work rose by about as much. The penalty on PhD-intensive titles shrank from $-0.24$ to $-0.07$, mostly before 2023. Yet in every year the same firm advertised an Application job in India more readily than a Building job. This is what the model predicts when distance becomes cheaper for all skilled work at once, and the largest jumps came in 2020–21, when work went remote. Nor does the gap wear off as a firm settles in: firms new to India show the same gap in their first year as several years later.

Generative AI changed which tasks firms hire for more than where they place them. Inside the same firm and quarter, U.S. postings for software developers who code applications fell by about 15 percent relative to general tech after ChatGPT's release, with no pre-trend, while postings for developers who build models and data systems show no lasting decline. Application postings in other occupations fell by nearly as much, so the decline follows the task rather than the occupation, and an occupation-level comparison would have missed which software developers it reached. A separate and briefer adjustment appeared at the bottom of the U.S. ladder. Measured against other technical occupations as in Westby et al. (2026), junior software-developer postings in the firms' U.S. offices fell by 19 percent relative to senior ones in 2023 and recovered by 2024, while in their Indian offices the estimate is close to zero.

Where firms place the tasks did not detectably change, and the model explains why. A technology that makes distant work easier to direct and check lowers the cost of distance and should shift tasks toward the cheaper locations, starting with the codified tasks near the margin. A technology that does part of the work itself, wherever the work is done, saves labor in every location and should leave the division where it was. Around the release of ChatGPT the data look like the second case. Relative to the same firm's general tech, no category's Indian share jumped by a statistically significant amount when compared with seven earlier placebo dates, and technical work as a whole shows no detectable shift against the firms' non-technical postings. The intervals are wide, allowing a rise of up to 41 percent, but the test is not blind: it picks up the large shifts of 2020–21, when work went remote. The one category that rose faster than at any placebo date is governance and security, the auditing of models and systems rather than their building.

U.S. pay kept pace throughout. As India's share of these tasks' postings doubled, the mean wage in the occupations of the U.S. postings rose about 26 percent between 2019 and 2025, as fast as other white-collar pay and as prices. This is the one-sided prediction of the model: when the productivity effect of placing tasks in other countries is at least as large as the labor-supply effect, U.S. pay in the occupations of these tasks does not fall behind

(Grossman and Rossi-Hansberg 2008).

The paper contributes to work on tasks and technology in labor markets, to the evidence on generative AI and early-career workers, and to the empirical study of trade in tasks; Section 2 reviews each. What it adds is a view of the task inside the same firm in five countries. One occupation code holds tasks that the same firm places in different countries, and the same title at the same firm differs in about a quarter of its task content from one country to another, so the task read from the job description, not the occupation or the title, is the unit at which firms divide technical work. Seen at that unit, the single ranking of tasks that trade theory assumes can be tested directly, and fails; and the change in U.S. hiring after ChatGPT can be traced to the software developers who code applications, separately from a brief dip in junior postings.

Section 2 describes the setting and the related literature, Section 3 the postings and Section 4 the task measure. Section 5 sets out the model, Section 6 shows where firms place each task and how India climbed the stack, and Section 7 what happened to U.S. pay. Section 8 turns to generative AI, what it changed in U.S. hiring and whether it moved the work, and Section 9 concludes.

# 2 Related literature

## 2.1 Technical work across countries

Large employers now run their technical work through their own offices and service centres in several countries and through IT-services firms that do the work for them. The firms studied here are of this kind: banks such as JPMorgan Chase and Citigroup, technology companies such as Oracle and Microsoft, and IT-services firms such as Accenture, IBM and Wipro, all of which advertise technical jobs in the United States and India. India is the largest destination. It grew into a software exporter in the 1990s (Arora et al. 2001), firms later began to place engineering and research work there as well as routine services (Lewin et al. 2009), and it now hosts more than 1,700 "global capability centres" of foreign companies, employing about 1.9 million people (Nasscom and Zinnov 2024). In the firms' own postings, India's share of their combined U.S. and Indian technical jobs rose from 21.6 percent in 2019 to 44.7 percent in 2025, and among the 510 firms observed in both countries in every year, whose coverage changed less, from 30 to 52 percent. The work is software, data pipelines, testing and analytics, not call centres. India's rise overlapped with the shift to remote work during the COVID-19 pandemic, which showed that more than a third of U.S. jobs could be done entirely from home (Dingel and Neiman 2020) and lowered the cost of directing work from

a distance. In these firms' postings, India's odds of receiving both Building and Application work jumped relative to general tech in 2020 and 2021 (Section 8.1), and the penalty on PhD-intensive titles shrank mostly before 2023 (Section 6.4). The timing is consistent with a fall in the cost of distance common to skilled work; it is not a causal estimate, and part of India's rise reflects LinkUp's expanding coverage of Indian career sites.

The other three countries give the firms different options. The Philippines is, like India, distant and low-wage and shares few working hours with the United States. Mexico is low-wage, near and in U.S. time zones. The United Kingdom pays close to U.S. wages and serves a large market of its own. Doctoral training among AI workers is far more common in the United Kingdom and the United States than in India or the Philippines (Pandit and Nambiar 2026), a difference the model of Section 5 turns into a prediction about where firms place graduate-intensive tasks.

Generative AI reached these firms' technical work quickly. ChatGPT, built on OpenAI's GPT-3.5 model, was released in November 2022, and language models can write and review code, draft documents and summarize meetings: tasks that technical workers in every location perform. The question for this paper is whether such a technology changed which tasks the firms hire for, where they place them, or both.

## 2.2 Tasks, technology and workers

The task approach describes a job by what the worker does rather than by the occupation's label. Computers substitute for routine tasks that follow explicit rules and complement the problem-solving and communication tasks of the people around them (Autor et al. 2003), and the resulting shifts in demand across tasks account for much of the change in the structure of employment and pay (Autor and Dorn 2013; Acemoglu and Autor 2011; Acemoglu and Restrepo 2022). Tasks vary substantially among workers within the same occupation and predict their wages (Autor and Handel 2013). A new technology displaces workers from the tasks it automates while new tasks reinstate labor elsewhere (Acemoglu and Restrepo 2019; Autor 2015), and new work appears where innovations complement what an occupation produces (Autor et al. 2024). In Europe, routine-biased technical change and the movement of tasks across borders together explain much of the polarization of employment (Goos et al. 2014), and in U.S. manufacturing tasks are sorted by their complexity among native workers, immigrants and workers in other countries (Ottaviano et al. 2013).

Job postings have become the standard way to see these tasks below the occupation. Requirements written into vacancies reveal how firms change the skills they ask for (Hershbein and Kahn 2018; Deming and Kahn 2018), the text of newspaper advertisements shows

that much of the long-run change in tasks happened within narrowly defined titles (Atalay et al. 2020), and Shukla (2026) shows that most of the tasks a posting advertises lie outside its occupation's O*NET task list. This paper reads the task from the duties each posting lists (Section 4) and takes the measure to the same firms' jobs in five countries, where one occupation code, and often one title, holds tasks that the firm places in different countries.

## 2.3 Generative AI and early-career work

Exposure measures predict which tasks AI and language models can do (Felten et al. 2021; Eloundou et al. 2024), and experiments and deployments show that generative AI raises the productivity of the workers who use it (Noy and Zhang 2023; Brynjolfsson et al. 2025b). Establishments whose tasks were exposed to an earlier wave of AI hired more AI workers and fewer workers in other positions (Acemoglu et al. 2022). Since ChatGPT, studies of U.S. vacancies, payroll records and résumés find weaker demand for early-career workers in exposed occupations, software developers above all (Westby et al. 2026; Brynjolfsson et al. 2025a; Hosseini and Lichtinger 2025), while in Denmark chatbot adoption has so far had small effects on earnings and hours (Humlum and Vestergaard 2025). Chandar and Klein Teeselink (2026) compare labor demand across 41 countries, Ravanilla (2026) finds slower employment growth in the more exposed parts of Philippine business-process outsourcing, through reduced hiring rather than layoffs, and Friebel et al. (2026) show how such a shock can reshape a firm's ladder of junior and senior workers, temporarily if it raises productivity and permanently if it speeds learning.

These studies observe one country at a time and measure exposure by occupation. None sees where the same firm does the work, so none can ask whether the technology changed where tasks are placed, and none splits an occupation into the tasks inside it. Section 8 does both: it tests for a shift in the firms' placement of tasks at the release of ChatGPT, reproduces the junior-hiring estimate of Westby et al. (2026) in the firms' U.S. offices and asks the same of their Indian offices, and uses the task measure to show which software-developer postings fell.

## 2.4 Trade in tasks

The tools for asking where a firm places each task come from the theory of trade in tasks. In Grossman and Rossi-Hansberg (2008), firms trade tasks rather than finished goods, and a task is placed in another country when the wage saving outweighs the cost of doing it at a distance; because one index of that cost ranks all tasks, every destination takes them in the same order (Grossman and Rossi-Hansberg 2012; Rodríguez-Clare 2010). Work on knowl-

edge hierarchies traces the cost to communication between the workers who solve routine problems and those who solve hard ones (Garicano 2000; Antràs et al. 2006; Antràs and Rossi-Hansberg 2009), and a neighbouring line assigns skill-intensive tasks to skill-abundant countries (Costinot 2009; Costinot and Vogel 2010), with Costinot et al. (2012) showing how to recover comparative advantage from a double comparison of countries and industries.

The empirical evidence has stayed a level above the task. It measures how easily an occupation's work can be done from another country (Blinder 2009; Blinder and Krueger 2013) and estimates the effects of moving work out of a country by industry, imported input or local labor market (Feenstra and Hanson 1996; Amiti and Wei 2009; Ebenstein et al. 2014; Crinò 2010; Liu and Trefler 2019; Autor et al. 2013; Hummels et al. 2014), reviewed by Hummels et al. (2018). The closest evidence on tasks is Becker et al. (2013), who find that German multinationals that place more work in low-income countries keep more non-routine and interactive tasks in Germany. Knowledge that must be communicated directly is costly to move (Keller and Yeaple 2013), firms whose work relies on complex communication serve foreign markets by exporting rather than producing there (Oldenski 2012), and firms choose among many source countries at once (Antràs et al. 2017). Section 5 extends Grossman and Rossi-Hansberg (2008) to several destinations, skill and job-level placement, which yields a gravity equation for jobs inside the firm, the job-level analogue of the trade workhorse (Head and Mayer 2014), and Section 6 tests the single ranking directly. The work that firms place in the United States most readily is of the skill- and information-intensive kind whose growth has concentrated in the largest U.S. cities (Eckert et al. 2022).

## 2.5 Contribution

The paper contributes to the labor literature on tasks and technology by adding a measure of the task read from the duties each job description lists, taken to the same firms' jobs in five countries; at that level one occupation code, and often one title, holds tasks that the same firm places in different countries, which is why the division of technical work has been invisible in occupation data (Section 4). To the evidence on generative AI and early-career work literature, it adds the cross-country side of the same firms: around ChatGPT's release the firms' placement of tasks across countries shows no detectable shift, while U.S. postings for software developers who code applications fell, a decline that follows the task rather than the occupation and is separate from a brief dip in junior postings (Section 8). To the study of trade in tasks, it adds a direct test of the single ranking of tasks inside the firm, which the data reject, and evidence that doctoral skill accounts for part of how firms divide the work among destinations (Sections 5 and 6).

# 3 Data

## 3.1 Postings collected from employers

The data are job postings collected by LinkUp, which scrapes employers' own career sites rather than aggregating job boards. The text of each posting is therefore the employer's own, with no aggregator duplicates, recruiter rewrites or listings re-posted by third parties, and every measure in this paper is read from that text.

The method has a cost. How many postings LinkUp records for a firm depends on how completely it scrapes the firm's career site in a given month as well as on how much the firm is hiring, and LinkUp's coverage of India expanded over the sample. For this reason no result interprets the absolute volume of postings. The estimators count advertised jobs, but they identify only relative composition: firm × country × period effects absorb how many postings LinkUp records for a firm in each country and period, and firm × task × period effects absorb how much of each task the firm advertises overall. The identifying assumption is that coverage does not vary differentially across task categories within a firm, country and period. The scrape itself is global: of the 60,207 employers with a located posting in 2019–2025, 48,678 post in the United States, 12,106 in the United Kingdom, 5,665 in India, 3,444 in Mexico and 2,198 in the Philippines (Appendix Figure 19). LinkUp does not record where a firm is headquartered.

## 3.2 Firms and countries

The firms are the 4,552 employers in LinkUp's U.S.–India technical-posting frame for 2019–2025. Nearly all advertise in both countries in some occupation (4,534 of the 4,544 found in LinkUp's multicountry extract), and 3,558 advertise technical jobs in both countries within 2019–2025; 806 advertise technical jobs only in the United States and 188 only in India in that window. I exclude KPMG, Deloitte and Infosys, which post several Indian jobs for every U.S. job, in proportions unlike any other firm, and are the subject of a companion paper. That leaves 4,549 firms. Including the three in the U.S.–India comparison leaves India's Building–Application gap at −0.20 (s.e. 0.04) and moves no destination's gap by more than 0.001 (Appendix Table 19). Of the 4,549, 47 have no labelled distinct job and are absent from the principal estimates: 33 never advertise a technical job with a full description, and 14 lose all their labelled postings when the title and creation date are matched from the multicountry extract. The principal sample therefore has 4,502 firms (Appendix Table 19). The four destinations are the firms' largest whose postings are written in English, so that all five countries can be labelled on identical terms, and they differ in ways the model of

Section 5 cares about. India and the Philippines are distant, low-wage and share few working hours with the United States. Mexico is low-wage, near and in U.S. time zones. The United Kingdom pays close to U.S. wages and serves a large market of its own.

The firms are not a balanced panel. The number with any posting in both the United States and India in the same year rises from 1,090 in 2019 to 2,048 in 2025, mostly through entry on the Indian side. Only 510 firms post in both countries in every year, but they hold 76 percent of the firms' Indian postings and 46 percent of their U.S. ones. Grouping firms under their listed parent with LinkUp's dated ticker reference, so that LinkedIn and GitHub count with Microsoft and Waymo and DeepMind with Google, gives 3,999 parents, 521 of them in both countries every year. Hiring is not dominated by a few firms: no firm holds more than 2 percent of the jobs in lookup titles, and the hundred largest hold 48.9 percent (Appendix Figure 20). Rebuilding the firm set directly from LinkUp's raw records finds 313 further parents with postings in both countries. They hold 0.5 percent of the relevant Indian postings, mostly from 2024–25, and adding them leaves the U.S.–India placement coefficients unchanged to the second decimal.

## 3.3 From postings to distinct jobs

Keeping the earliest copy of any record that appears in more than one of LinkUp's 291 raw shards, the firms have 55,968,263 postings in 2019–2025: 45.4 million in the United States, 5.87 million in India, 3.43 million in the United Kingdom, 867,043 in Mexico and 441,002 in the Philippines.

A posting is not yet a job. LinkUp records a job once for every location listed on it, so a job advertised in five U.S. cities appears five times. Copies are more common in the United States than in India and differ by kind of work (Appendix G), so counting records would overstate the U.S. share of work advertised in many cities. The unit of analysis is instead the *distinct job*: one firm, one country, one title and one creation date, the day LinkUp first recorded the posting. Titles are lower-cased and every run of characters other than letters and digits becomes a space, so "Sr. Software Engineer – Java" and "sr software-engineer (java)" are the same title; synonyms such as "senior" and "sr" are not merged. A firm that advertises a role in both the United States and India makes two placements, which is what the paper counts. The rule can err either way, since two openings with the same title posted on the same day count as one, and a title that names its city does not collapse with its copies. Collapsing copies leaves 30.6 million distinct jobs across all titles.

### 3.4 The analysis sample and what is labelled

The analysis sample is every technical job of the firms: distinct jobs whose LinkUp O*NET code is among the 48 technical codes of Appendix L. Each is labelled from the full text of its posting (Section 4), so a job enters only if its posting has a description in English. In the United States and India, 72 percent of the 13.6 million technical postings have a full description; in the other three countries 95.7 to 97.7 percent do, and at least 90 percent in every country and year. Postings are English in all but 0.1 to 0.2 percent of cases everywhere except Mexico, where 85.4 percent are (Section 4.5 gives the test). A distinct job takes the label of its earliest posting. The sample has 6,089,734 distinct jobs: 3,645,133 in the United States, 1,765,046 in India, 467,728 in the United Kingdom, 120,531 in Mexico and 91,296 in the Philippines. India's share of the firms' combined U.S. and Indian technical jobs rose from 21.6 percent in 2019 to 44.7 percent in 2025.

Skill intensity (Section 4.6) and location cost (Appendix M) are measured for titles rather than for jobs. They use a fixed list of titles, the title lookup of Appendix H, which covers 1,148,667 of the distinct jobs (Appendix Tables 21 and 20).

### 3.5 Required experience

For the analysis of job levels I parse the minimum years of experience each posting requires. A number counts only when it appears within forty characters of the word "experience," so that "a three-year degree" does not, and a distinct job keeps the lowest requirement across its copies. The rungs are fixed before estimation: *junior* (0–2 years), *mid* (3–5) and *senior* (6 or more), with the split of Westby et al. (2026) at four years reported for comparison. A requirement is stated in 60 percent of U.S. and 58 percent of Indian technical postings, 48 percent of Philippine, 43 percent of Mexican and only 17 percent of British ones. Indian postings ask for more experience in almost every occupation: over 2019–2025, 21 percent of Indian software-developer postings are junior against 37 percent of U.S. ones (Appendix Tables 22 and 23). Appendix D asks where the gap comes from.

## 4 Seven task categories read from posting text

The unit of this paper is a task, and I measure it from what a posting says the worker will do. This section explains why that is necessary, how the measure is built, how well it does, and what it shows before any regression.

## 4.1 Why occupation codes cannot measure task location

The natural unit for a study of where work is done would be the occupation. Every U.S. and Indian posting in LinkUp carries an O*NET occupation code matched from its title (Appendix L), but for this question the code is both too coarse and too fixed. Most of the tasks a posting advertises lie outside its occupation's O*NET task list (Shukla 2026).

It is coarse because one code, Software Developers (15-1252), contains 39.2 percent of India's technical postings and 25.4 percent of the United States'. India's distribution across codes is much more concentrated (Herfindahl index 0.185 against 0.112), so whatever distinguishes the technical work these firms do in India from the work they do in the United States mostly happens inside a single code.

It is fixed because the code is assigned from the title, and the same title can describe very different work. Figure 1 shows four postings live on two employers' careers sites in October 2026. The first firm's two "Data Scientist II" jobs share a title, a level and an occupation code. In San Diego the worker invents prototypes in named-entity recognition, pricing, agentic AI and large language models and takes them into production. In Bengaluru the worker designs experiments and causal models and also owns "the data pipelines, metrics, and reporting infrastructure," including the ETL and the dashboards. The first is model building; the second is functional application with a share of data infrastructure. The U.S. job asks for two years of experience, the Indian one for four. A second large technology firm's two mid-level data scientists make the same point with the requirement held fixed: both ask for five years and a quantitative degree, yet the San Bruno job joins a team building LLM-powered search for a video platform while the Bengaluru job builds alerting systems and self-service analytics tools and maintains "foundational data pipelines and single-source-of-truth dashboards."

The same split appears inside the LinkUp data. In 2024 JPMorgan Chase advertised two "Data Scientist" jobs in the job family it calls Predictive Science, one in Jersey City and one in India (Appendix Table 26). The first asks the worker to develop machine-learning models and the pipelines that train and serve them; the second to write SQL, build Tableau dashboards and automate reporting. The occupation code and the title record them as the same job. The full-text measure reads the first as model building and the second as data and ML infrastructure.

The examples should not be over-read. Across all "data scientist" postings in 2024–25, about half are classified as Building in each country (48 percent in the United States, 51 percent in India), and the shares are nearly identical within the 359 firms that post the title in both. The title does not mean one thing in the United States and another in India. Rather, a single title, at a single firm in a single year, covers both kinds of work, and only

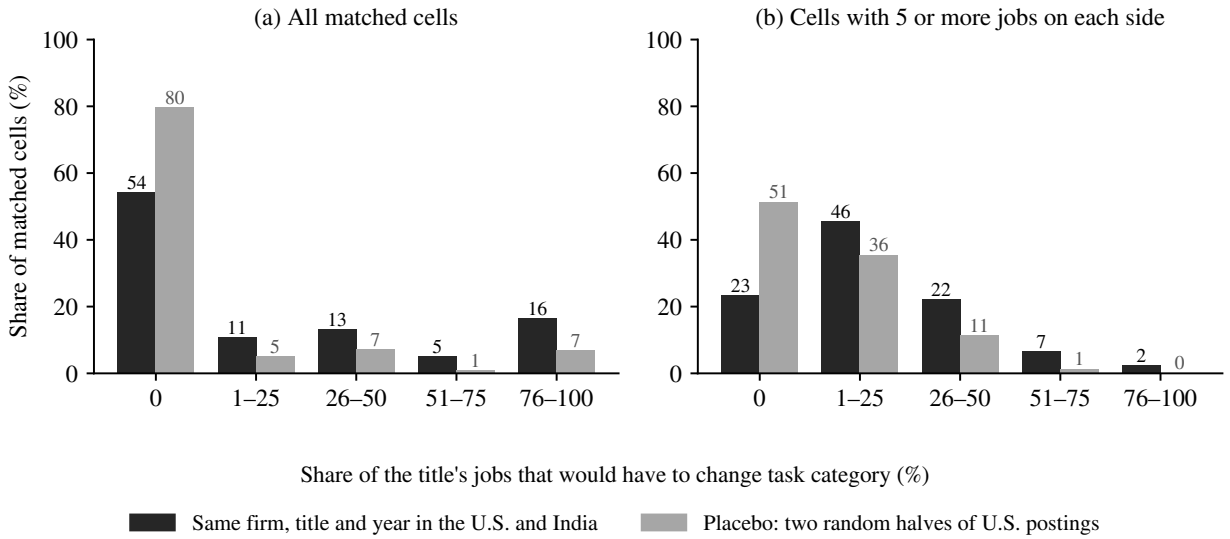



Figure 2: Same title, different work: how much of the task mix differs between the U.S. and Indian versions of the same firm, title and year. For each firm × title × year with distinct jobs in both countries, the share of its jobs that would have to change full-text task category to make the two countries' task mixes identical. Dark bars: U.S.–India cells. Light bars: the same statistic between two random halves of U.S. postings, reweighted to the same cell sizes. Panel (a): all 58,940 matched cells; in a cell with one job on a side the share is 0 or 100 percent. Panel (b): the 5,991 cells with at least five jobs on each side. English postings, 2019–2025; KPMG, Deloitte and Infosys excluded.

the text can tell them apart. The point holds across the whole sample (Figure 2). Among firm × title × year cells with at least two jobs, 30 percent of U.S. and 43 percent of Indian jobs sit in cells that span more than one category. In the 58,940 cells where a firm advertises the same title in the same year in both countries, the two countries' most common category differs in 25.8 percent and the most common group (Building, Application, general tech) in 15.2 percent, against 11.6 and 7.0 percent between two random halves of U.S. postings of the same size; the disagreement exceeds the placebo in every cell-size bin. Part of the mixing within a country is classification error, so these shares bound the real differences in work from above, but the excess over the placebo is not error of that kind.

The difference is just as clear on the whole task mix. To give the U.S. and Indian versions of the same firm, title and year the same mix of tasks, 26.7 percent of their jobs would have to change category, against 11.5 percent between random halves of U.S. postings of the same size; the share is 23.2 percent for the Philippines, 24.2 for Mexico and 18.0 for the United Kingdom, against about 11.5 in each case (Appendix Table 28). Figure 2 shows how the difference is spread. In 54 percent of U.S.–India cells the two task mixes are identical, against 80 percent between random halves of U.S. postings, and in 16 percent at least three-quarters of the jobs would have to change, against 7 percent. Among the cells with five or more jobs on each side, where the share can take many values, the two mixes are identical in 23 percent of cells, against 51 percent by chance. The difference does not, however, run systematically out of Building. Within the same firm, title and year, a job posted in India is 2.4 percentage points (s.e. 0.6) more likely to be Application and 2.2 points (s.e. 0.4) less likely to be general tech than its U.S. twin, and the Philippines and Mexico look the same, but the Building share of the same title is statistically indistinguishable from zero in every destination (Appendix Figure 24). The opening example therefore shows that one title can hold different work in two countries, not that the U.S. version is typically the Building work and the other country's version the Application work. Inside a single title the work is mixed in both countries: among "software engineer" jobs that share their firm and year with another posting of the title, 62 percent sit in firm-title-years whose postings span more than one category, and half span more than one of Building, Application and general tech (62 and 49 percent in India). The text, not the title, is what separates the software engineers who build models and data systems from those who code applications, the distinction on which Section 8.2 rests. The Building–Application gaps of Section 6 come from which titles and jobs firms place in each country, which is why the paper measures the task of each job rather than of each title.

## 4.2 The seven categories

The four postings in Figure 1 sit in different parts of one ecosystem. Behind the AI search product the San Bruno team is building, someone trains and evaluates the models; someone builds the pipelines, feature stores and serving systems that feed and run them; someone writes the application code that calls them; someone designs and tests the product; someone reports what it does for the business; someone secures it; and someone keeps the firm's networks and servers running. Each is a separate job with its own skills, and each can be done in a different country. The first firm's San Diego job sits at the first step, the two Bengaluru jobs mostly at the reporting step and in the pipelines. The categories follow these steps, one sentence each:

- *T1 Model building*: develops and trains predictive and machine-learning models, designs architectures, prepares training data, runs experiments, evaluates accuracy;
- *T2 Data and ML infrastructure*: data pipelines and warehouses, feature stores, model deployment and serving, monitoring, distributed and cloud data infrastructure;
- *T3 Technical application*: develops applications and services, connects systems through APIs, writes, tests and maintains code;
- *T4 Product and testing*: product management, interface and interaction design, user research, customer-facing front ends;
- *T5 Functional application*: enterprise and business systems, reporting and dashboards, business analysis and process improvement;
- *T6 Governance and security*: information-security controls, audits, risk frameworks, regulatory compliance and policy. Someone who looks over the features of a model and evaluates them for risk, as a model hub such as Hugging Face does before hosting it, does this work: the job is about models, but its duties are review and risk assessment, not training;
- *T7 General tech*: networks, servers, systems administration, cloud operations, help desk and desktop support.

The first two make up *Building*, the work of producing the systems; the next four make up *Application*, the work of using, testing and governing systems others build; general tech, the work of keeping the firm's IT running, is the benchmark against which every other category is measured. The split follows the engineering account of production machine-learning systems,

in which the code that trains a model is a small part of a deployed system and data collection, feature pipelines, serving and monitoring make up most of the effort (Sculley et al. 2015): T1 is the first, T2 the second. An analogy from manufacturing: Building is designing the engine and the plant that makes it; Application is fitting the engine into cars, deciding where it is used, testing the car and certifying it; general tech keeps the factory's machines running.

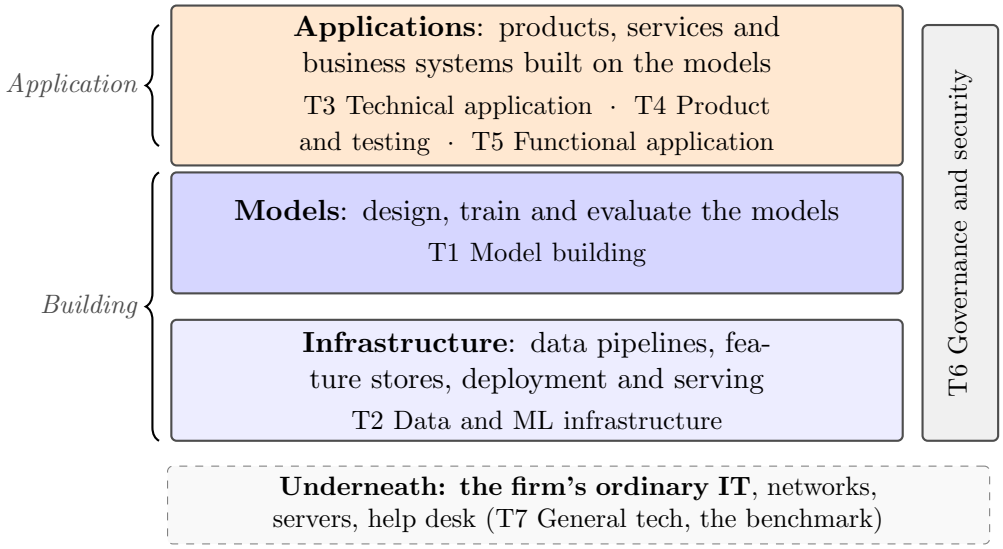



Figure 3: The three tiers of the technology ecosystem and the seven task categories. The tiers follow the industry's description of the generative-AI stack (Bornstein et al. 2023); the split of Building (T1, T2) into models and infrastructure follows Sculley et al. (2015). Application (T3–T6) is the top tier together with governance and security; general tech (T7), the firm's ordinary IT, is the benchmark.

The categories also follow the industry's own description of the generative-AI ecosystem as three tiers: infrastructure at the bottom, models in the middle and applications on top (Bornstein et al. 2023). Most of the large firms in this paper work in all three at once: a bank runs its own data platform, trains some models and builds many applications on them, and an IT-services firm does all three for its clients. Figure 3 places the seven categories in the tiers. Governance and security (T6) runs alongside all three, because every tier has to be secured and audited, and it is counted with Application because its work is reviewing systems rather than producing them.

Table 1 lists the categories with typical jobs and their weight in each country. The two stacks differ in shape: technical application is India's largest category, 37 percent of its technical postings in 2025 against 14 percent in the United States, while U.S. postings are spread more evenly and hold proportionally more model building and governance and security. Building is a larger share of the U.S. stack in every year, and it grew in both countries after 2022: from 16.3 to 22.5 percent of U.S. technical postings between 2022 and

Table 1: The seven task categories and their weight in U.S. and Indian technical postings

| | | United States | | India | |
|---|---|---|---|---|---|
| Category | Typical jobs | 2019 | 2025 | 2019 | 2025 |
| *Building* | | | | | |
| T1 Model building | machine-learning engineer, applied scientist, data scientist who builds models | 8.4 | 10.9 | 5.5 | 6.8 |
| T2 Data and ML infrastructure | data engineer, cloud data engineer, MLOps engineer | 8.1 | 11.5 | 6.2 | 9.9 |
| *Application* | | | | | |
| T3 Technical application | software or application engineer, systems engineer, embedded engineer | 16.3 | 13.6 | 33.1 | 36.6 |
| T4 Product and testing | product manager, UX designer, QA or test engineer | 14.6 | 14.7 | 9.5 | 9.4 |
| T5 Functional application | business or data analyst, ERP and CRM consultant, project manager | 17.1 | 13.0 | 18.5 | 14.8 |
| T6 Governance and security | security analyst, information-security engineer, risk and compliance analyst | 15.5 | 17.7 | 7.9 | 7.9 |
| *Benchmark* | | | | | |
| T7 General tech | IT support specialist, system or network administrator, help-desk technician | 20.0 | 18.6 | 19.3 | 14.6 |

*Notes:* Percent of each country's technical postings at the panel firms in the year (columns sum to 100). Categories assigned from the full posting text (Section 4.3); KPMG, Deloitte and Infosys excluded. Typical jobs are frequent titles in each category.

2025, and from 13.3 to 16.7 percent of Indian ones.

## 4.3 Assigning postings to categories

The assignment uses a sentence-embedding model from natural-language processing. Such a model has been trained on a very large amount of text and turns any passage into a long list of numbers, so that passages that mean similar things get similar numbers. Picture the result as a map. Each posting becomes a point, and postings that describe similar work land close together even when they use different words: two postings about adapting large language models end up near each other whether one says "LLM fine-tuning" and the other "customizing foundation models," a match that no keyword list written in advance would make.

I put the seven category sentences on the same map and assign each posting to the sentence it lands closest to. The posting is read in full, its title together with its whole description, up to about 2,800 words. Nothing in this step is fitted to the outcomes studied later, and the same sentences and rule label postings in every country and year (Appendix I gives the steps and formulas).

Figure 4 shows the map for 60,000 postings, flattened to two dimensions. Model building and data infrastructure sit next to each other, with technical application beside them; functional application, product and testing and governance lie in another part, and general tech in the middle. The regions overlap, as they should, because many jobs mix duties: a data engineer who also trains models sits between the two. In Panel B, Building work forms clusters of its own rather than being scattered through Application. Drawn country by country (Figure 5), no country's postings occupy a region of their own, which is what allows one set of sentences to label all five.

The category sentences contain no AI vocabulary, so the measure cannot drift when employers start to mention AI, and it reaches back before generative AI existed. Reading the whole posting matters because U.S. postings are about twice as long as Indian ones: a fixed truncation would read a quarter of a U.S. posting and half of an Indian one. Reading the full text changes about a fifth of the labels, equally in the two countries.

## 4.4 Checking the labels

One check uses titles whose category is not in doubt (machine-learning engineer, help-desk technician, SAP consultant and the like): 5,063 such postings in a test corpus of 24,558. Accuracy is averaged across the seven categories, so that a large category cannot carry it. With the title in the input, the labels match the expected category for 90.5 percent of them,

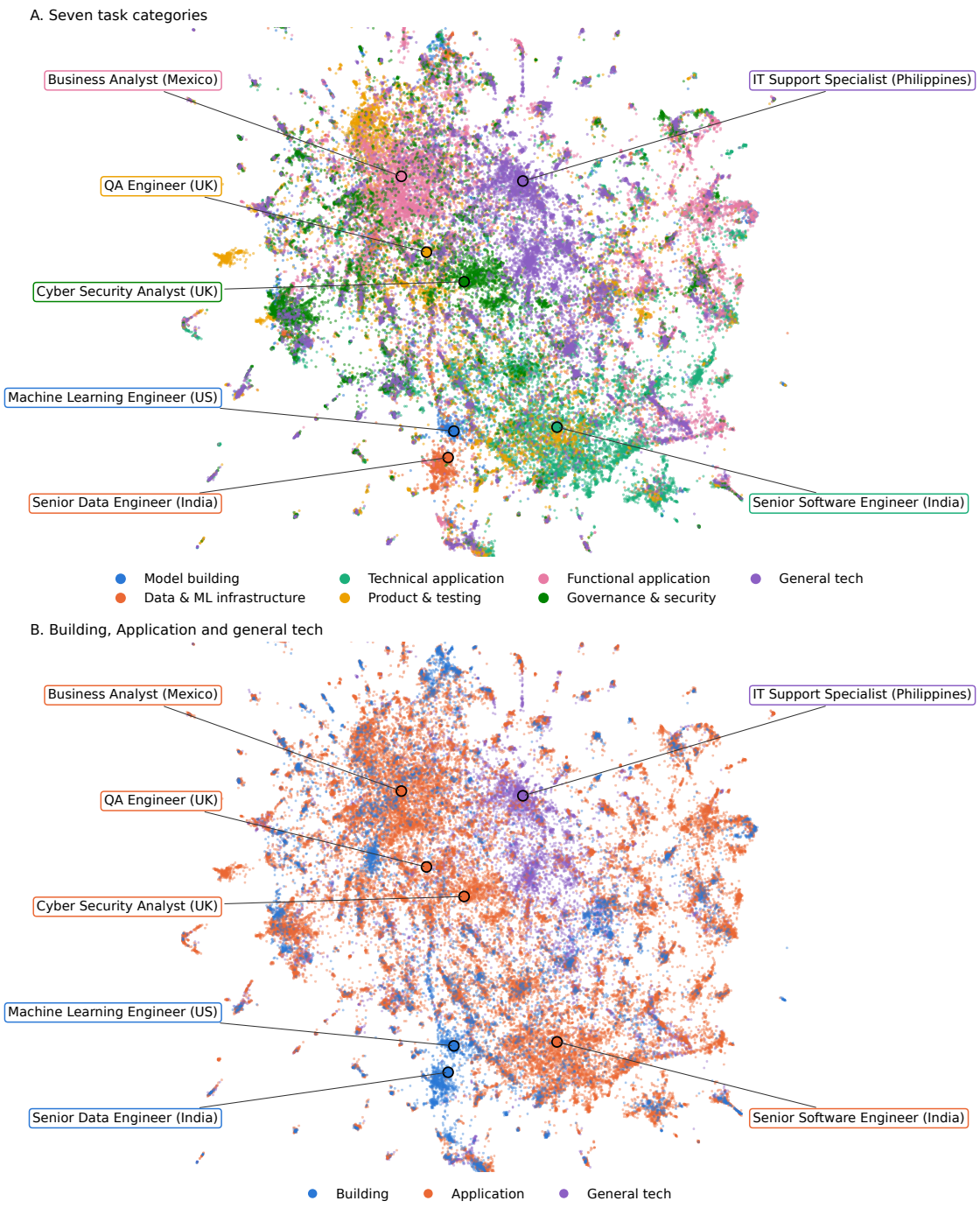



Figure 4: What the text measure sees. Each dot is one of 60,000 technical postings (12,000 from each of the five countries, 2019–2025), placed so that postings with similar duties are close together; the map is a two-dimensional projection (UMAP) of the posting embeddings and only a picture: it enters no estimate. Panel A: colour by the category each posting is assigned to; labels sit where each category is densest. Panel B: Building (model building, data and ML infrastructure), Application (the other four focal categories) and general tech. Outlined dots are single postings, labelled with their actual job title and country. The picture shows the central 94 percent of points on each axis.

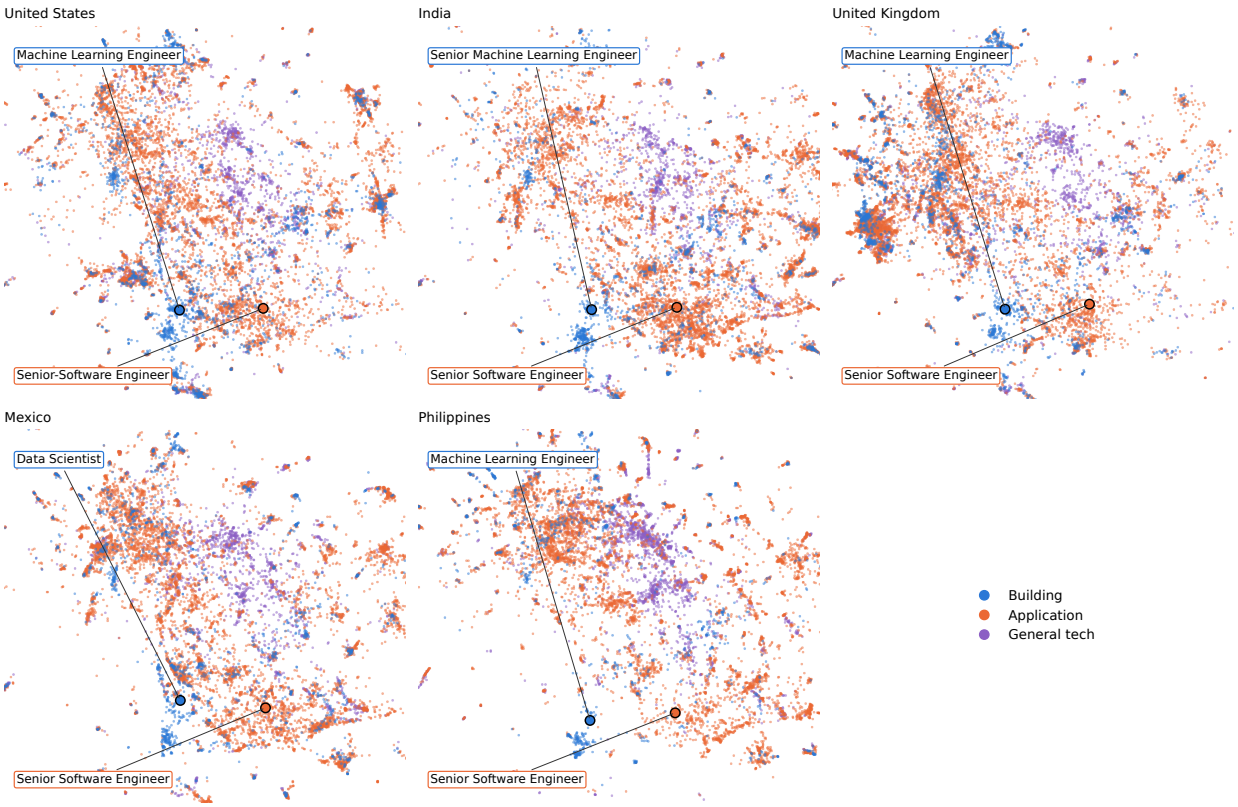



Figure 5: The same map, one panel per country: each country's own postings, coloured by group. In each panel the outlined dots are that country's postings nearest to where model building and technical application are densest, labelled with their actual job titles. Window as in Figure 4.

and U.S. and Indian accuracy are within 11 points of each other in every category. Because this is partly a check of the title, I repeat it twice. With the title removed from the input, accuracy falls to 73.9 percent, just under the 75 percent I set in advance as a pass mark. With the title removed and its phrase also deleted from the body, the hardest version, it is 59.7 percent. The text alone identifies model building (80 percent), governance and security (88), general tech (90) and product and testing (82) well, but confuses technical with functional application.

Title-blind, the U.S.–India accuracy gap reaches 19 to 20 points in three categories, so error that differs by country cannot be ruled out. Its direction can be measured, so I carry the errors into the within-firm estimates (Appendix K; Appendix I gives the simpler correction of pooled shares, which points the same way). For every distinct job I draw a category from the probability, given its label and its country, that it truly belongs to each category, computed by Bayes' rule from that country's title-blind confusion matrix, and I re-estimate Equation (2) on each of 40 draws. Because the title-blind matrices treat every job as if its title said nothing, they overstate the errors in the labels actually used, and the exercise is a stress test. Misclassification attenuates the gap. Corrected, India's Building–Application gap widens from $-0.20$ to $-0.39$, the Philippines' from $-0.46$ to $-0.86$ and Mexico's from $-0.30$ to $-0.57$, and it is negative and significant in every draw. The adversarial version, which gives U.S. jobs India's error pattern and Indian jobs the U.S. one, is the only one that shrinks India's gap, to $-0.17$, still significant in every draw. The six-category profile keeps its order (rank correlations of corrected with observed coefficients of 0.83 to 1.00 in the three low-wage destinations). The United Kingdom is the exception: its gap of $-0.01$ becomes $-0.19$ with its own matrix and $-0.18$ with India's, and stays near zero only with the U.S. matrix (0.00) or in the swap ($-0.05$), so its small gap is the least robust of the four.

The labels could still fail in particular places or at particular times, and the remaining checks look for that (Appendix K, Table 30). Accuracy did not change at the release: with the title it is 90.2 percent before November 2022 and 90.8 percent after in the United States, and 91.9 and 91.0 percent in India; title-blind, it rose slightly after the release (4.3 points, s.e. 1.6), equally in both countries (India minus the United States $-1.4$, s.e. 3.8). In the other three destinations, on 97 to 137 clear-title test postings each, accuracy with the title is 80.7 percent in the United Kingdom, 83.4 in Mexico and 81.6 in the Philippines, lower mainly because technical application is confused with functional application. Labels of the same role do not drift: when a firm posts the same title in the same country six months apart, the label is unchanged for 83 percent of U.S. and 81 percent of Indian roles, and across the release that persistence changes by $-1.1$ points in the United States (s.e. 1.0) and by $+2.1$ points more in India (s.e. 1.8). Inside ambiguous titles the labels predict an

independent signal: the named tools that a separate extraction model (GLiNER) finds in the same posting, from lists that leave out generic AI words. Inside the same firm's Software Developer titles, Building jobs name data and model tools 6.0 percentage points more often than Application jobs (s.e. 0.7, on a base of 3.2 percent), Application jobs name application languages and frameworks 4.3 points more often (s.e. 0.5, base 41 percent), and general-tech jobs name infrastructure tools 3.0 points more often (s.e. 0.4, base 2.0 percent). The Building gap is larger in India (+1.8, s.e. 0.8) and unchanged after the release (+0.6, s.e. 1.1), so the labels separate duties as well there and then.

The categories are not a relabelling of occupations. Appendix J compares them with job titles and O*NET codes: they agree with both where they should, and they vary within occupations. Inside Software Developers, U.S. postings moved toward Building: between 2019 and 2025 the share in model building rose from 7.0 to 10.4 percent and in data and ML infrastructure from 7.0 to 11.3, while technical application fell from 52.2 to 44.8.

## 4.5 Following the categories into five countries

The same seven sentences label every technical posting of the panel firms in the United Kingdom (593,934 postings), Mexico (172,869) and the Philippines (123,360), with the same embedding and settings (Equation (28)); U.S. and Indian postings keep their labels. The procedure reproduces those labels: relabelling the first 300,000 postings of the U.S.–India corpus returns the stored label for every one of them. The embedding model reads English. A posting counts as English when, among its first 300 words, at least 60 percent of the words on a fixed list of common English and Spanish function words are English. The primary estimates keep English postings; Section 6.2 adds the others.

Figure 6 shows what the measure finds before any regression. Building is 17.8 percent of U.S. technical jobs and 17.5 percent of British ones, against 16.0 percent in India, 14.2 in Mexico and 10.7 in the Philippines. Each destination also has a category of its own: technical application is 27.3 percent of India's technical jobs against 15.7 percent of U.S. ones; functional application is 25.6 percent of the Philippines'; product and testing and governance and security are each 19.7 percent of the United Kingdom's. The shares mix what firms place in each country with how much they hire there, which is why Section 6 compares them within firm and year.

India's rise was not confined to routine work. Its share of the firms' combined U.S. and Indian jobs rose from 23 to 47 percent in Application, from 17 to 43 percent in Building and from 21 to 40 percent in general tech between 2019 and 2025 (Figure 7, Panel A); part of every rise is LinkUp's expanding coverage of India (Section 3.1).

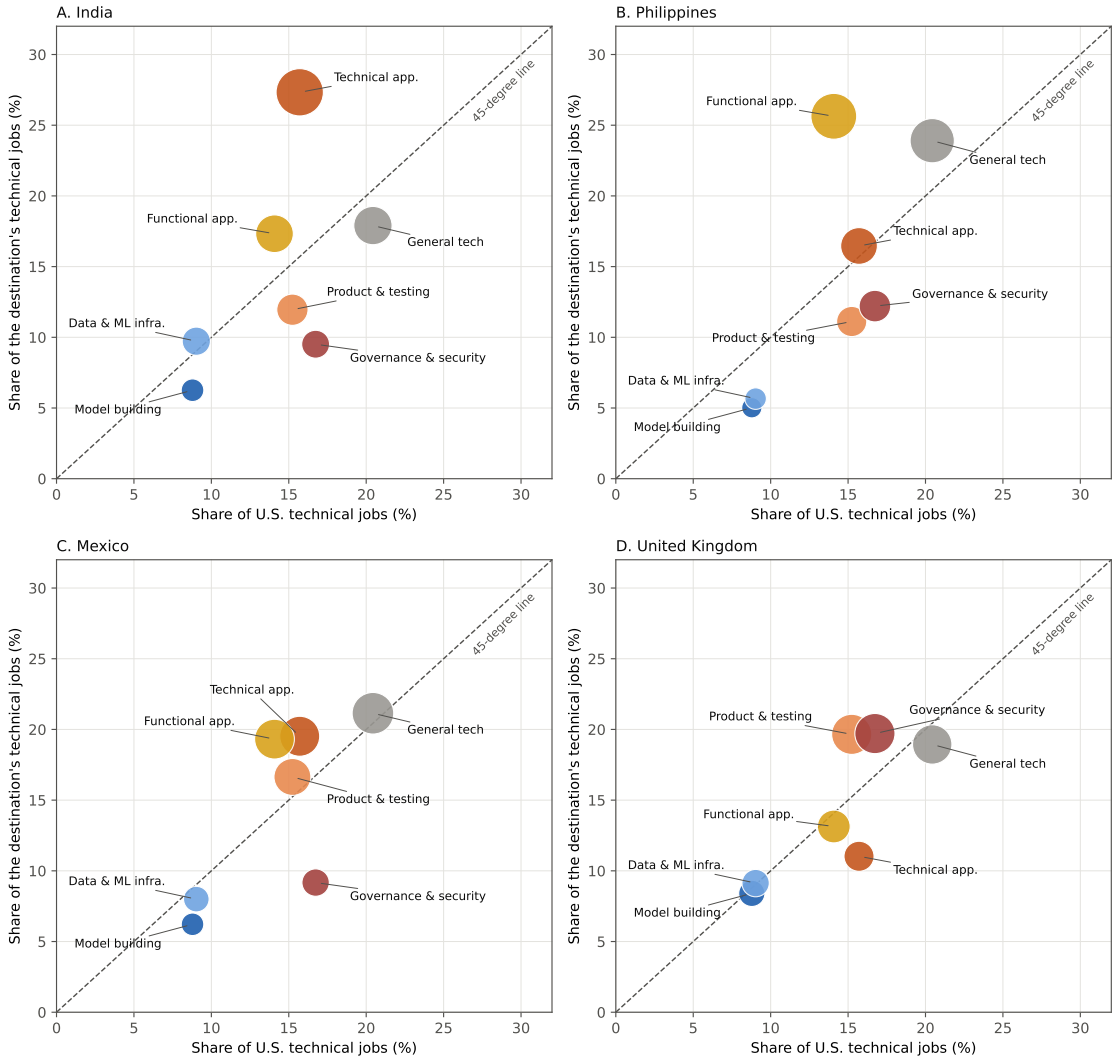



Figure 6: Each destination's task mix against the U.S. mix, 2019–2025. Each circle is a task category: the horizontal axis is its percent of the panel firms' U.S. technical jobs, the vertical axis its percent of the destination's technical jobs, and circle area is proportional to the destination's jobs in the category. Blue: Building; orange and red: Application; grey: general tech. A circle above the dashed 45-degree line is a larger share of the destination's work than of U.S. work. Distinct jobs (firm × country × title × creation date), English-language postings, full-text categories (Section 4.5); KPMG, Deloitte and Infosys excluded. Pooled shares mix what firms place in each country with which firms hire there and with LinkUp's coverage; the within-firm estimates are in Table 5.

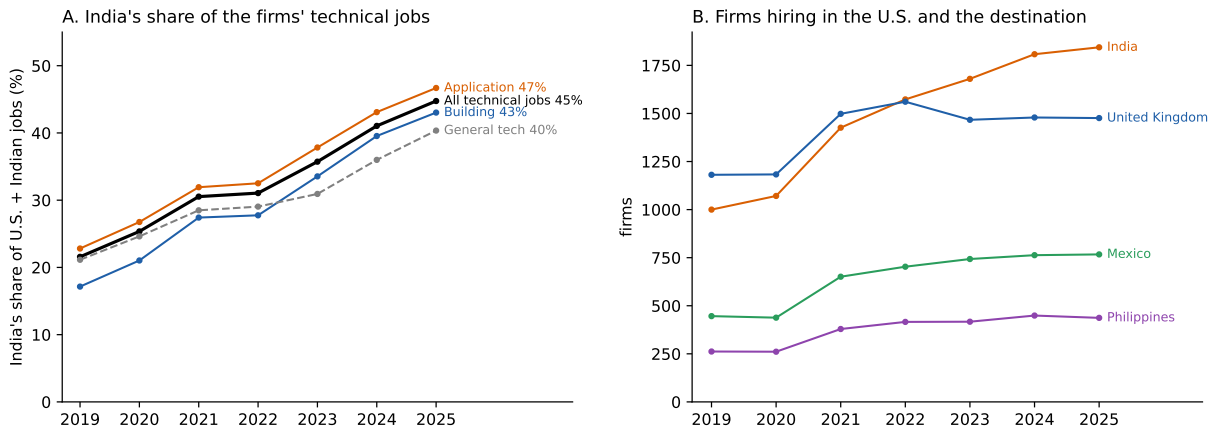



Figure 7: The analysis sample, 2019–2025. Panel A: India's share of the firms' combined U.S. and Indian technical jobs, all technical jobs and by group (Building: model building, data and ML infrastructure; Application: the other four focal categories; general tech). Panel B: firms with at least one technical job both in the United States and in the destination in the same year. Distinct jobs (firm × country × title × creation date), all technical O*NET codes, full-text categories; KPMG, Deloitte and Infosys excluded. Shares partly reflect LinkUp's expanding coverage of India and are not used as outcomes.

The measure also separates jobs that share an occupation code. Figure 8, Panel A, places the twelve largest occupation codes, which hold 84 percent of U.S. and Indian technical postings, across the seven categories. Only the support and security codes are mostly one kind of work. Computer and Information Systems Managers spread over five categories in both countries, and Data Scientists, almost three-quarters of whose postings are Building, apply analysis to a business function in a fifth of them. The same code also holds different work in the two countries, as Panel B shows for Software Developers in all five. Technical application is the largest category everywhere, but its share ranges from 45.8 percent in the United Kingdom and 48.5 percent in the United States to 65.4 percent in India and 69.7 percent in the Philippines, and the Building share runs the other way: 17.0 percent of U.S. software-developer postings and 15.3 percent of British ones, against 10.3 percent in India, 10.6 percent in Mexico and 7.5 percent in the Philippines. An occupation-level measure would treat all of these jobs as equally easy to move.

The title lookup of Appendix H assigns a category from the job title alone. It is not the paper's measure of tasks. It is kept because it is the fixed list of titles on which the title-level attributes (skill intensity, location cost and the wage comparisons) are measured, and because it gives a second set of estimates that does not depend on the text classifier (Table 25). It covers about a sixth of the technical postings in the three new countries. Table 2 compares the two labels on the same postings: rows are lookup categories, columns are full-text categories.

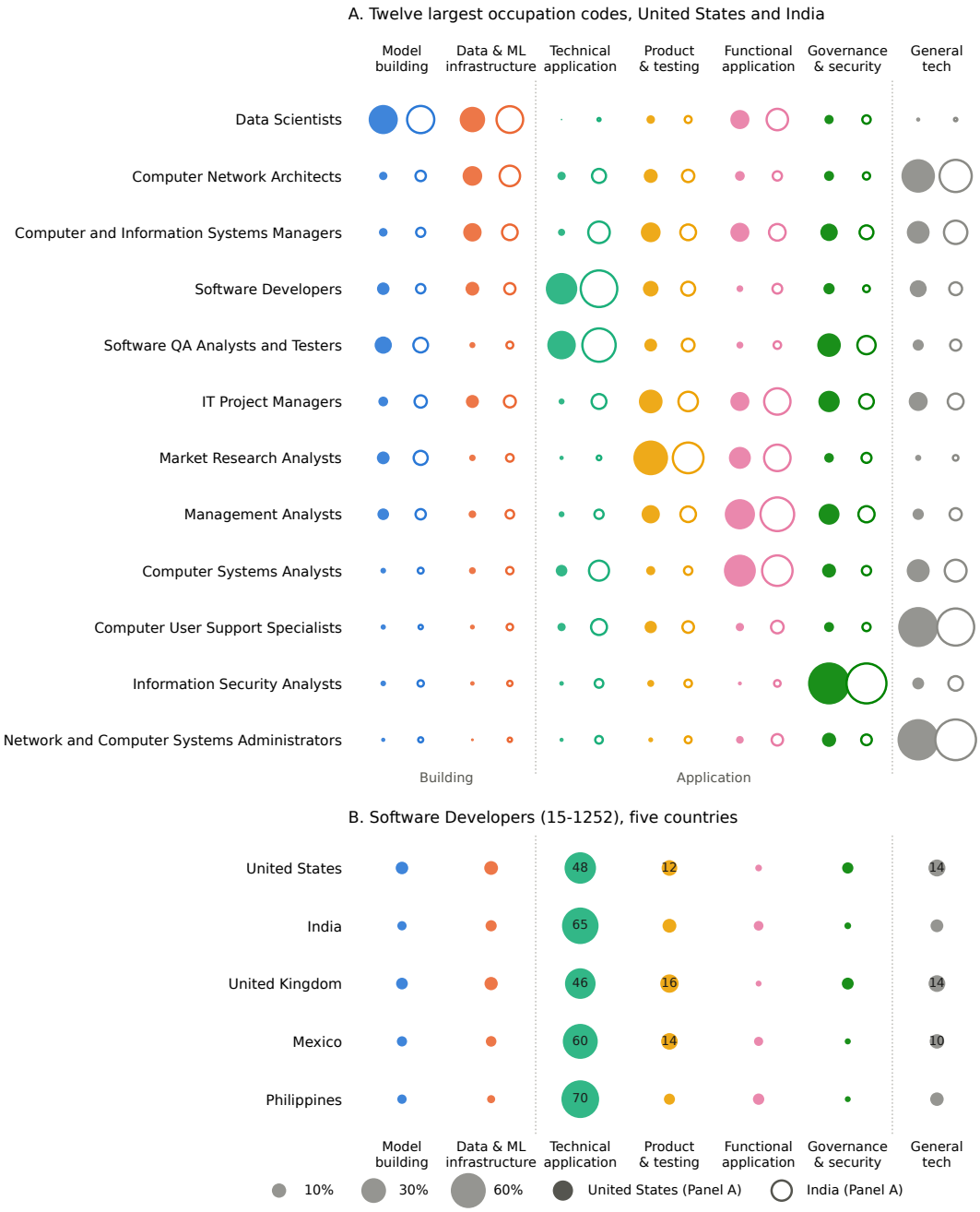



Figure 8: One occupation code, many tasks: full-text categories of postings by O*NET occupation, 2019–2025. Circle area is proportional to the percent of the occupation's postings in each category, on the same scale in both panels. Panel A: the twelve largest occupation codes by U.S. and Indian postings (84 percent of 9,097,904), filled circles for the United States and open circles for India; rows ordered by the U.S. Building share. Panel B: postings coded Software Developers (15-1252) in each country, with the percent written where it is 10 or more; 1,586,029 postings in the United States, 1,218,032 in India, 111,540 in the United Kingdom, 36,668 in Mexico and 21,462 in the Philippines. KPMG, Deloitte and Infosys excluded.

Table 2: Where the title lookup's categories go under the full-text label, by country (row percent)

| Rows: title-lookup category | | Columns: full-text category (main measure) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Country | T1 | T2 | T3 | T4 | T5 | T6 | T7 |
| T1 Model building | U.S. | **29.9** | 17.6 | 3.0 | 5.7 | 2.7 | 4.1 | 37.1 |
| | India | **59.2** | 4.6 | 0.9 | 2.2 | 0.5 | 0.9 | 31.7 |
| | U.K. | **35.4** | 8.1 | 2.1 | 3.5 | 0.5 | 1.9 | 48.5 |
| | Mexico | **31.7** | 9.7 | 0.4 | 1.3 | 3.6 | 1.1 | 52.1 |
| | Philippines | **15.2** | 2.5 | 0.2 | 0.5 | 0.7 | 2.7 | 78.1 |
| T2 Data and ML infrastructure | U.S. | 6.3 | **32.1** | 8.9 | 5.7 | 2.6 | 6.0 | 38.5 |
| | India | 3.7 | **29.4** | 27.5 | 2.0 | 1.1 | 1.8 | 34.6 |
| | U.K. | 15.0 | **27.9** | 8.7 | 2.5 | 1.0 | 4.5 | 40.4 |
| | Mexico | 4.0 | **29.4** | 14.2 | 1.0 | 1.6 | 1.4 | 48.4 |
| | Philippines | 1.4 | **23.9** | 21.4 | 0.5 | 1.4 | 2.7 | 48.7 |
| T3 Technical application | U.S. | 8.9 | 4.2 | **38.6** | 7.0 | 11.3 | 11.0 | 18.9 |
| | India | 3.5 | 2.6 | **70.1** | 5.8 | 4.4 | 2.4 | 11.1 |
| | U.K. | 6.7 | 3.2 | **46.8** | 8.0 | 2.0 | 7.8 | 25.6 |
| | Mexico | 3.4 | 1.1 | **59.2** | 12.0 | 4.3 | 2.5 | 17.6 |
| | Philippines | 3.9 | 1.5 | **63.6** | 3.4 | 3.0 | 3.1 | 21.5 |
| T4 Product and testing | U.S. | 9.9 | 5.4 | 36.0 | **22.9** | 5.3 | 11.5 | 9.0 |
| | India | 9.4 | 1.3 | 60.5 | **16.1** | 1.6 | 8.2 | 2.8 |
| | U.K. | 7.7 | 1.3 | 34.9 | **29.4** | 1.8 | 20.8 | 4.1 |
| | Mexico | 14.7 | 2.8 | 44.3 | **20.7** | 1.2 | 9.7 | 6.5 |
| | Philippines | 11.9 | 0.8 | 54.2 | **18.6** | 1.4 | 10.7 | 2.5 |
| T5 Functional application | U.S. | 5.2 | 4.8 | 2.7 | 22.4 | **31.9** | 21.8 | 11.3 |
| | India | 4.8 | 4.0 | 14.1 | 13.6 | **43.4** | 9.8 | 10.3 |
| | U.K. | 4.4 | 6.0 | 1.8 | 28.2 | **31.1** | 19.9 | 8.5 |
| | Mexico | 5.5 | 5.0 | 3.9 | 25.4 | **39.0** | 10.7 | 10.5 |
| | Philippines | 5.6 | 2.6 | 7.5 | 16.6 | **46.6** | 11.9 | 9.1 |
| T6 Governance and security | U.S. | 1.0 | 2.5 | 1.2 | 0.9 | 1.7 | **85.7** | 6.9 |
| | India | 0.6 | 0.5 | 0.8 | 0.7 | 0.7 | **92.9** | 3.8 |
| | U.K. | 0.8 | 0.4 | 0.4 | 1.1 | 0.9 | **92.5** | 3.9 |
| | Mexico | 1.5 | 0.2 | 1.7 | 0.9 | 2.2 | **88.9** | 4.6 |
| | Philippines | 0.4 | 0.1 | 0.0 | 0.1 | 2.3 | **94.9** | 2.2 |
| T7 General tech | U.S. | 2.6 | 6.7 | 6.5 | 11.8 | 8.5 | 23.8 | **40.0** |
| | India | 0.5 | 1.7 | 12.7 | 2.7 | 3.3 | 6.8 | **72.4** |
| | U.K. | 3.0 | 1.1 | 3.0 | 5.7 | 1.2 | 6.0 | **80.0** |
| | Mexico | 0.0 | 0.3 | 3.2 | 2.0 | 1.3 | 0.8 | **92.2** |
| | Philippines | 0.0 | 0.2 | 2.1 | 2.1 | 1.9 | 1.7 | **92.1** |

*Notes:* Rows are the fixed title lookup (Appendix H), which assigns a category from the job title alone; columns are the full-text category of Equation (28), the paper's main measure: T1 model building, T2 data and ML infrastructure, T3 technical application, T4 product and testing, T5 functional application, T6 governance and security, T7 general tech. Each row gives the percentage of the lookup category's postings in the country that fall in each full-text category, and sums to 100. Bold: the full-text category with the same name. The two measures are built differently, so the bold cells are not expected to be near 100; the off-diagonal cells show where a title misstates the duties in the text, chiefly network-engineering titles in lookup model building. English-language postings of the panel firms in lookup titles, 2019–2025: 2,048,368 in the United States, 456,116 in India, 107,879 in the United Kingdom, 26,085 in Mexico and 22,140 in the Philippines.

Because the two are built differently, the table is not an accuracy check; its off-diagonal cells show where a title misstates the duties its posting describes. Governance titles stay in governance and security (86 to 95 percent in every country). Technical application titles go mostly to technical application, 39 percent in the United States and 47 to 70 percent elsewhere. The large change is in lookup *model building*. Only 30 percent of its U.S. postings are model building in the full text, 59 percent in India, 35 in the United Kingdom, 32 in Mexico and 15 in the Philippines; most of the rest is general tech (37, 32, 48, 52 and 78 percent), the network engineering that the title lookup grouped with model training. The full-text categories therefore remove network work from Building. Lookup general tech titles are general tech in 72 to 92 percent of postings outside the United States but in only 40 percent of U.S. postings, where a quarter are security work. Two rows move for a different reason. Product and testing titles are product and testing in only 16 to 29 percent of postings and technical application in 35 to 60 percent: the full text reads many quality-assurance and test postings as software engineering. Functional application titles are functional application in 31 to 47 percent of postings, with most of the rest in product and testing (14 to 28 percent) and governance and security (10 to 22 percent).

## 4.6 The skill intensity of a task

The model's second task attribute is skill intensity, $s_z$: how much of a task's labor must come from workers with graduate training. I measure it from what U.S. employers ask for. For every title in the lookup, including the ordinary technical titles, I compute from 2.34 million U.S. postings with full text the share that ask for a PhD or doctorate. Secondary measures record the share that ask for a master's degree and the share that ask for a research record (publications, peer-reviewed or first-author work); the exact terms are in Appendix G. Each share is standardized across titles, weighting by U.S. postings. Across all titles, 2.8 percent of U.S. postings ask for a PhD and 12.5 percent for a master's degree; one standard deviation of the PhD share is 12.5 percentage points.

The measure orders the categories as one would expect. Averaged over U.S. jobs in each full-text category, the title's PhD share is 14.1 percent in model building, 5.2 percent in technical application and 2.8 percent in data and ML infrastructure, against 2.5 percent in governance and security, 1.8 percent in general tech, 1.3 percent in product and testing and 1.2 percent in functional application. Among large titles, the highest shares belong to applied scientists, machine-learning engineers and biostatisticians. Skill intensity is unrelated to the measured location cost of Appendix M (posting-weighted correlation across titles $-0.01$), so the two measures capture different reasons a task might stay in the United States.

The measure has a known flaw. Some employers' templates list an education-substitution ladder ("Bachelor's and eight years, Master's and six, PhD and four"), so a mention of a PhD is not always a requirement. A few titles reach a PhD share near one this way, and the ladders are common in U.S. government-contractor postings. Section 6.3 shows that the main result survives dropping those titles, the high-clearance titles and small titles, and that the measure is informative at the top of the distribution of PhD shares.

# 5 How firms divide the stack: the model in brief

The technology ecosystem of Section 4.2 is a chain of jobs, and a firm decides where each link is done. Take a bank that builds an AI assistant for its customers. Training the model is model building (T1); feeding it account data is data and ML infrastructure (T2); the phone app is technical application (T3); designing and testing the conversation is product and testing (T4); the business dashboards are functional application (T5); reviewing the model for risk before release is governance and security (T6); keeping the servers running is general tech (T7). The bank does not decide where to make the assistant; it decides where to put each of these seven jobs. That observation—firms trade tasks, not finished products—is the starting point of Grossman and Rossi-Hansberg (2008), henceforth GRH. This section gives the intuition of the extension I estimate and the predictions it makes. Each formal statement appears where it is tested, and the full model, its assumptions and the proofs are in Appendix A.

A job of task $z$ placed in country $c$ costs

$$C_{ijzc} \;=\; \frac{a(z)\,\tau_c(z)\left(w_c^H\right)^{s_z}\left(w_c^L\right)^{1-s_z}\phi_{ic}}{\varepsilon_{ijc}}, \tag{1}$$

the labor it needs, $a(z)$, times three prices: the extra cost of doing the task far from the firm's U.S. teams, $\tau_c(z)$, which is one in the United States; the wage, a mix of graduate and non-graduate pay weighted by the task's graduate share $s_z$; and the firm's own cost of operating in the country, $\phi_{ic}$. The last term, $\varepsilon_{ijc}$, stands for the reasons invisible to the researcher that make one job cheaper in one place, such as a team that needs someone in its time zone. Each job goes where it is cheapest. Put numbers on the bank. Suppose each job costs \$100 in the United States and \$30 in wages in India. The phone app is easy to specify and check, so done from Bengaluru it costs \$45 and it moves there. Training the model needs the bank's customer data, compute and daily contact with its research team, so from Bengaluru it would cost \$120, and it stays in New York.

Three features make this GRH's model of trade in tasks extended to technical work. Each

destination has its own cost of distance for each task, so London and Manila need not find the same jobs easy. Tasks differ in how much graduate labor they need and countries in how much of it they have, so PhD-heavy work is expensive to place where PhDs are scarce. And because the job-level draws spread the jobs of one task across countries, the share of a task placed in each country takes the form of a gravity equation, the workhorse of international trade. Think of the bank as a buyer of tasks and of each country as a seller. The number of model-training jobs the bank places in India depends on how many such jobs it needs, on how attractive India is for all of its work—wages there, the size of its Bengaluru office, how fully LinkUp sees its Indian postings—and on how costly model training in particular is to do from India. Fixed effects absorb the first two, which is why the firm's wages, office size and LinkUp coverage cancel, and the estimates isolate the third (Proposition 1).

## 5.1 What the model predicts

Table 3: What the model predicts and where each prediction is tested

| Prediction | In words | Tested in |
|---|---|---|
| P1. Selective placement | Every destination receives some categories more and others less than general tech. | Table 5 |
| P2. Own rankings | Destinations order the categories differently; one common ranking would make every pair agree. | Section 6.2, Appendix M |
| P3. What stays in the United States | Building moves out of the United States less than Application, most of all where graduates are scarcest. | Tables 5 and 7 |
| P4. Sorting on skill | PhD-heavy titles stay in the United States, more so where PhDs are scarcer. | Table 7 |
| P5. Climbing the stack | A destination that gets cheaper gains every category at once; its Building gap closes only if its graduate supply or Building's own cost changes. | Figure 12 |
| P6. Generative AI | Labor-saving AI moves no task: no jump in India's shares at ChatGPT's release; hiring may fall wherever the task is done. | Tables 9 and 14 |
| P6′. The ladder | A short junior dip in the United States; a shift toward mid-level jobs in India. | Table 15 |
| P7. U.S. pay | U.S. pay in the work that moved does not fall behind comparable work. | Section 7 |
| P7′. One labor market | Pay in Building and Application rises alike. | Section 7 |

*Notes:* Formal statements: Propositions 1–5 above and 7–8 in Appendix B. Table 13 in the same appendix gives the model object behind each estimate.

The predictions are about which jobs move, not how many. India's lower wages and

the cost of the bank's Bengaluru office are the same for the app and for model training, so they decide how much of its work the bank places in India but not which jobs. Which jobs move is decided by the two costs that differ across tasks, the cost of distance and the cost of graduate skill. Table 3 lists the predictions and where each is tested; in words they are these.

The work that needs scarce graduates and close contact with U.S. teams, Building, stays in the United States more than the work of applying systems others build, and most of all where graduates are scarcest (P1, P3). Because each destination prices distance in its own way, the destinations need not rank the tasks in the same order, as a single cost of distance would force them to (P2). Titles that ask for a PhD stay in the United States, and more so the scarcer PhDs are in the destination (P4). A destination that becomes cheaper for all skilled work at once gains every task together, and its Building gap closes only if its graduate supply rises or Building's own cost of distance falls (P5).

A new technology can act on either part of the cost. If it makes distance cheaper—it writes the documentation, reviews the code, summarizes the meeting—it lowers $\tau$ and moves work out of the United States, as remote work did. If it instead does part of the work itself in every location, it lowers $a(z)$ everywhere. A coding assistant that cuts the phone app's labor by 30 percent in both countries makes it cost \$70 in the United States and \$31.50 in India: India is still cheaper by exactly the same ratio, so the job does not move, and the bank simply advertises fewer such jobs in both places (P6). Finally, moving the app to Bengaluru saves the bank money, which works like a rise in productivity for its U.S. workers in the remaining tasks, while the workers whose tasks moved compete for other jobs. If the first force wins, U.S. pay in the work that moved does not fall behind comparable work (P7).

## 5.2 From the model to the estimates

The model becomes a regression once the task category stands for the task. With the category $k$ as the task and $V_{ickt}$ the number of distinct jobs (Section 3.3) of firm $i$ in country $c$, category $k$ and year $t$, Proposition 1 becomes

$$\mathbb{E}[V_{ickt}] = \exp\Big(\alpha_{ikt} + \gamma_{ict} + \sum_{d \in \mathcal{D}} \sum_{k \neq g} \mathbb{1}[c = d]\,\mathbb{1}[k]\; b_{d,k}\Big), \tag{2}$$

estimated by Poisson pseudo-maximum likelihood, with general tech ($g$) as the reference category. In plain terms, $b_{d,k}$ is a log odds ratio: the odds that a category-$k$ job goes to $d$ rather than the United States, divided by the same odds for a general-tech job. A coefficient of 0.22 means odds $e^{0.22} = 1.25$ times as high, 25 percent higher; a coefficient of $-0.15$ means

odds 14 percent lower. Grouping the categories into Building $B = \{\text{T1}, \text{T2}\}$ and Application $A = \{\text{T3}, \ldots, \text{T6}\}$ gives the comparison at the center of the paper,

$$\Delta_d \equiv b_{d,B} - b_{d,A}, \qquad \widehat{\text{se}}(\Delta_d) = \sqrt{\hat{V}_{BB} + \hat{V}_{AA} - 2\hat{V}_{BA}}, \tag{3}$$

with $\hat{V}$ the firm-clustered covariance matrix of the two coefficients. $\Delta_d$ is negative when the firm places Building work in $d$ less than Application work, and $1 - e^{\Delta_d}$ is the shortfall in Building's odds. In the model each coefficient is $-\theta$ times the destination's extra cost of the category, part location cost and part the cost of its graduate labor (Appendix Equation (15)).

GRH's single ranking also becomes a statement the data can reject. If every destination ranked tasks by one index $t_k$, scaled only by its distance $\lambda_d$, every pair of destinations would order the categories identically:

$$b_{d,k} = -\theta\, \lambda_d\, (t_k - t_g) \quad \Longrightarrow \quad \text{corr}_{\text{rank}}\big(b_{d,\cdot},\, b_{d',\cdot}\big) = 1 \text{ for all } d, d'. \tag{4}$$

I report the rank correlation of the six $b_{d,k}$ for each pair of destinations; a negative correlation rejects the single ranking.

# 6 Which work stays in the United States

## 6.1 The design

Table 4: A worked example: one firm's technical jobs in one year

| | United States | India | India's share of the firm's jobs (percent) | Odds, India / United States |
|---|---|---|---|---|
| Building | 100 | 40 | 28.6 | 0.40 |
| Application | 300 | 240 | 44.4 | 0.80 |
| General tech | 100 | 50 | 33.3 | 0.50 |
| All technical jobs | 500 | 330 | 39.8 | 0.66 |

*Notes:* Hypothetical numbers. Category coefficients: $\log(0.40/0.50) = -0.22$ for Building and $\log(0.80/0.50) = +0.47$ for Application. Building–Application gap: $\log(0.40/0.80) = -0.69$, equal to the log of Building jobs per Application job in India (40/240) over the same ratio in the United States (100/300).

When a firm advertises a given kind of technical job, what are the odds that it advertises it in India, the Philippines, Mexico or the United Kingdom rather than in the United States, and do those odds differ across kinds of work? The comparison is always inside one firm in one year. A job's location is the country in which it is advertised; the data do not observe

whether it is filled or where the hired worker sits. Table 4 works it through for a firm that in 2024 has 500 technical jobs in the United States and 330 in India.

Take each kind of work in turn and divide the firm's jobs in India by its jobs in the United States. These odds are 0.4 for Building, 0.8 for Application and 0.5 for general tech. A category's coefficient compares its odds with those of general tech, so it is $\log(0.4/0.5) = -0.22$ for Building and $\log(0.8/0.5) = +0.47$ for Application. The Building–Application gap compares the two odds directly, and general tech drops out: $\log(0.4/0.8) = -0.69$, so a Building job's odds of being placed in India rather than the United States are half those of an Application job. The same number emerges if one looks inside each country instead. The firm posts one Building job for every six Application jobs in India and one for every three in the United States, and $\log\big((1/6)/(1/3)\big) = -0.69$. The gap is thus neither a share of the firm's postings nor a share within one country; it compares the firm's mix of tasks in India with its mix in the United States.

This double comparison is what protects the design against the obvious concerns. Whatever leads the firm to place much or little of its work in India in a given year—its wages, the size of its Bengaluru office, how well LinkUp scrapes its Indian career site—scales the entire India column by a common factor and therefore cancels from these ratios. The same is true of anything that makes the firm hire more of one kind of work in both countries, which scales a row. What remains is how the firm divides each kind of work between the two countries. The estimates average such comparisons over thousands of firms and seven years, giving more weight to larger firm-years.

Figure 9 draws this comparison for every firm with enough jobs. Each circle is a firm, placed by the log of its Building jobs per Application job in the United States on the horizontal axis and in the destination on the vertical axis, so that its vertical distance from the 45-degree line is the firm's own gap. The firm × country × year effect leaves no trace in the picture, since it shifts both of a firm's counts in a country by the same proportion. Firms differ a great deal. In each low-wage destination only a little over half of them lie below the line, but the largest firms do, which is why 70 to 81 percent of the jobs lie below it. Weighted by jobs, the firm gaps average $-0.19$ in India, $-0.41$ in the Philippines, $-0.29$ in Mexico and $+0.02$ in the United Kingdom, close to the estimates of Table 5.

The estimating equation is the gravity equation of Section 5.2, written out with every term labelled. For firm $i$, country $c \in \{\text{US}, \text{IN}, \text{PH}, \text{MX}, \text{UK}\}$, task category $k \in \{\text{T1}, \ldots, \text{T7}\}$

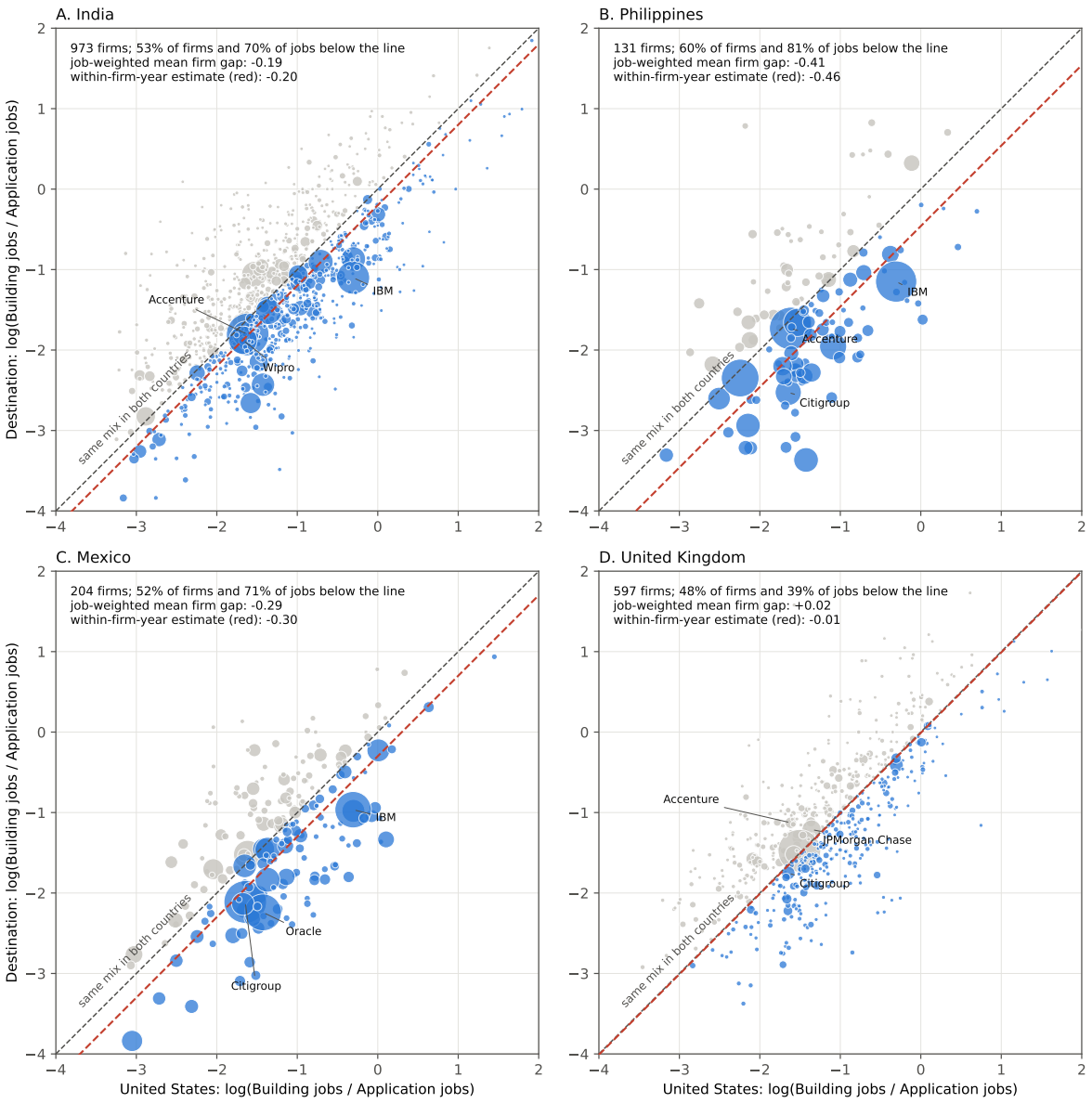



Figure 9: The Building–Application gap firm by firm. Each circle is a firm with at least 10 Building and 10 Application jobs in both the United States and the destination, 2019–2025 pooled (India 973 firms, the Philippines 131, Mexico 204, the United Kingdom 597); circle area is the firm's Building and Application jobs in the destination. Axes: log of Building jobs per Application job. Grey dashed line: the same mix in both countries (gap zero); blue circles lie below it. Red dashed line: the within-firm-year estimate of Table 5, Panel B. Distinct jobs, English-language postings, full-text categories; KPMG, Deloitte and Infosys excluded.

Table 5: Task categories relative to general tech, by destination: all technical jobs, full-text categories

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| *A. Category, relative to general tech (T7)* | | | | |
| T1 Model building | −0.07 | −0.59*** | −0.36*** | −0.06 |
| | (0.05) | (0.11) | (0.08) | (0.04) |
| T2 Data and ML infrastructure | 0.10** | −0.58*** | −0.28*** | −0.08* |
| | (0.05) | (0.10) | (0.07) | (0.05) |
| T3 Technical application | 0.72*** | −0.10 | 0.17* | −0.25*** |
| | (0.05) | (0.17) | (0.09) | (0.05) |
| T4 Product and testing | −0.10* | −0.52*** | −0.13** | 0.15*** |
| | (0.05) | (0.14) | (0.06) | (0.04) |
| T5 Functional application | 0.24*** | 0.22** | 0.10 | −0.10** |
| | (0.06) | (0.10) | (0.06) | (0.04) |
| T6 Governance and security | −0.36*** | −0.38*** | −0.37*** | −0.17*** |
| | (0.07) | (0.11) | (0.09) | (0.05) |
| *B. Grouped (separate regression)* | | | | |
| Building (T1–T2) | 0.04 | −0.59*** | −0.32*** | −0.08* |
| | (0.04) | (0.09) | (0.06) | (0.04) |
| Application (T3–T6) | 0.24*** | −0.13 | −0.02 | −0.07* |
| | (0.04) | (0.09) | (0.05) | (0.03) |
| Building − Application | −0.20*** | −0.46*** | −0.30*** | −0.01 |
| | (0.04) | (0.05) | (0.04) | (0.03) |
| *C. Building − Application, other samples* | | | | |
| All languages | −0.20*** | −0.46*** | −0.33*** | −0.01 |
| | (0.04) | (0.05) | (0.04) | (0.03) |
| Lookup-title jobs only | −0.15*** | −0.42*** | −0.11* | −0.09* |
| | (0.06) | (0.09) | (0.07) | (0.05) |
| Title-lookup labels, title cells | −0.10 | −0.53*** | −0.39*** | −0.14* |
| | (0.10) | (0.15) | (0.11) | (0.08) |
| Rank correlation with India | – | 0.43 | 0.83 | −0.43 |

*Notes:* PPML of Equation (2): counts of distinct jobs (firm × country × normalized title × creation date) in firm × country × category × year cells, 2019–2025, zeros filled for all five countries and seven categories in every firm-year with a job. Every technical posting of the panel firms (LinkUp O*NET code among the 48 of Appendix L) with a description, labelled from its full text. Fixed effects firm × category × year and firm × country × year. Panels A and B: English-language postings; 6,089,734 distinct jobs, 4,502 firms (the 47 other panel firms have no labelled distinct job; Appendix Table 19), 762,265 cells, of which 282,703 remain after Poisson drops cells whose fixed-effect groups contain only zeros; 3,692 firms have at least one year with jobs in both the United States and a destination, the firm-years that identify the coefficients. Building − Application and its standard error from the grouped regression's covariance matrix. Panel C: the same estimate including non-English postings; on the jobs in lookup titles only (full-text labels); and on title-lookup labels in firm × title × year cells (Table 25). Rank correlation: Spearman, six categories. KPMG, Deloitte and Infosys excluded. Standard errors clustered by firm in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

and year $t \in \{2019, \ldots, 2025\}$,

$$\underbrace{V_{ickt}}_{\substack{\text{jobs of task } k \\ \text{firm } i \text{ places in } c}} = \exp\Bigg( \underbrace{\alpha_{ikt}}_{\substack{\text{firm's need for} \\ \text{task } k \text{ anywhere}}} + \underbrace{\gamma_{ict}}_{\substack{\text{country's pull for} \\ \text{all the firm's work}}} + \sum_{d \in \mathcal{D}} \sum_{k' \neq \text{T7}} \underbrace{b_{d,k'}}_{\substack{\text{how } d \text{ prices} \\ \text{task } k'}} \mathbb{1}[c = d]\mathbb{1}[k = k'] \Bigg) \cdot \eta_{ickt} \tag{2}$$

with $\mathbb{E}[\eta_{ickt} \mid \alpha, \gamma] = 1$. In trade terms the firm is the buyer, the country is the seller and the task is the product. Each component has a plain meaning.

- $V_{ickt}$, the outcome, is a count: the number of distinct jobs (Section 3.3) of category $k$ that firm $i$ places in country $c$ in year $t$. The unit of observation is this cell. A firm that posts in a year contributes a cell for all five countries and all seven categories, including cells with no jobs. Keeping those zeros matters, because a firm that never places model building in Mexico is informative; Poisson pseudo-maximum likelihood (PPML) keeps them, a regression in logs would drop them (Santos Silva and Tenreyro 2006).

- $\exp(\cdot)$ and $\eta_{ickt}$. The count is modelled as the exponential of the terms in brackets times a mean-one error, as in trade gravity. PPML needs only that the mean is right, not that the counts are Poisson.

- $\alpha_{ikt}$, one for each firm, category and year, is the buyer's demand for the product: how many jobs of task $k$ the firm needs that year, wherever they are done. A firm that expands its data team in 2023 raises its data-infrastructure jobs in every country; this term takes that out.

- $\gamma_{ict}$, one for each firm, country and year, is the seller's general competitiveness for this buyer: everything about the firm's presence in the country that year—wages there, the size of its office, the general cost of working with that office and how fully LinkUp scrapes its career site. These are the terms that cancel in Table 4, and this is why the growth of LinkUp's coverage of India cannot drive the results.

- $\mathbb{1}[c = d]\mathbb{1}[k = k']$ switches on in the cells of destination $d$ and category $k'$ only.

- $b_{d,k}$, the coefficient of interest, is how destination $d$ prices task $k$ for these firms, the log odds ratio of Section 5.2. The United States and general tech (T7) are the baselines, which is why they have no coefficient; there are $4 \times 6 = 24$ coefficients.

The identifying assumption is that anything specific to a firm's presence in a country—including LinkUp's coverage—may vary by firm, country and year but does not favor some of the firm's tasks over others there. There are no other controls, and none are needed: anything that varies by firm, country and year, or by firm, category and year, is in the fixed effects. What the design cannot absorb is something that varies by firm, country and category at once, such as a firm opening a dedicated model-training lab in India in one year. Standard errors are clustered by firm, so all of a firm's cells, in every country and year, are allowed to be correlated. A firm-year with no jobs in a destination says nothing about placement there and drops out, so each destination's coefficients come from firms that post both there and in the United States in the same year. Each firm-year contributes a small table of countries by categories; with both sets of fixed effects the estimator is consistent as the number of jobs per firm-year grows (on bias in gravity models with several sets of fixed effects, see Weidner and Zylkin 2021), and because Poisson weights cells by their size, large firm-years carry the estimates.

The Building–Application gap of Equation (3) is the same regression with the six focal categories grouped into two. Attributes that belong to a job title rather than a job—PhD intensity (Section 6.3), measured location costs (Appendix M) and the title-lookup categories (Appendix H)—use the same equation with titles $z$ in place of categories,

$$\mathbb{E}[V_{iczt}] = \exp\Big(\alpha_{izt} + \gamma_{ict} + \sum_{d} \mathbb{1}[c = d]\; x_z'\beta_d\Big), \tag{5}$$

where $x_z$ is the title's attribute and $\alpha_{izt}$ a firm × title × year effect.

Every coefficient compares one destination with the United States. A comparison between two destinations comes from the same regression, because the United States cancels: $b_{IN,k} - b_{MX,k}$ is the log odds of India rather than Mexico. With four destinations the design cannot say which feature of a destination—distance, time zone, language or the wage gap—drives its pattern, and the estimates describe how firms divide work in equilibrium rather than the effect of a shock. I read each destination through the profile set out in Section 3.2 instead.

## 6.2 Which categories stay and which move (P1–P3)

The model's prediction for this comparison is Proposition 3 (Appendix A). Differencing the log odds of a Building and an Application task in the same destination removes the destination's wages and office costs and leaves two parts, $\Delta_d = \mathcal{L}_d + \mathcal{S}_d$: a location part, the extra cost of doing Building far from the firm's U.S. teams, and a skill part, Building's larger graduate share times the destination's graduate premium over the United States. The

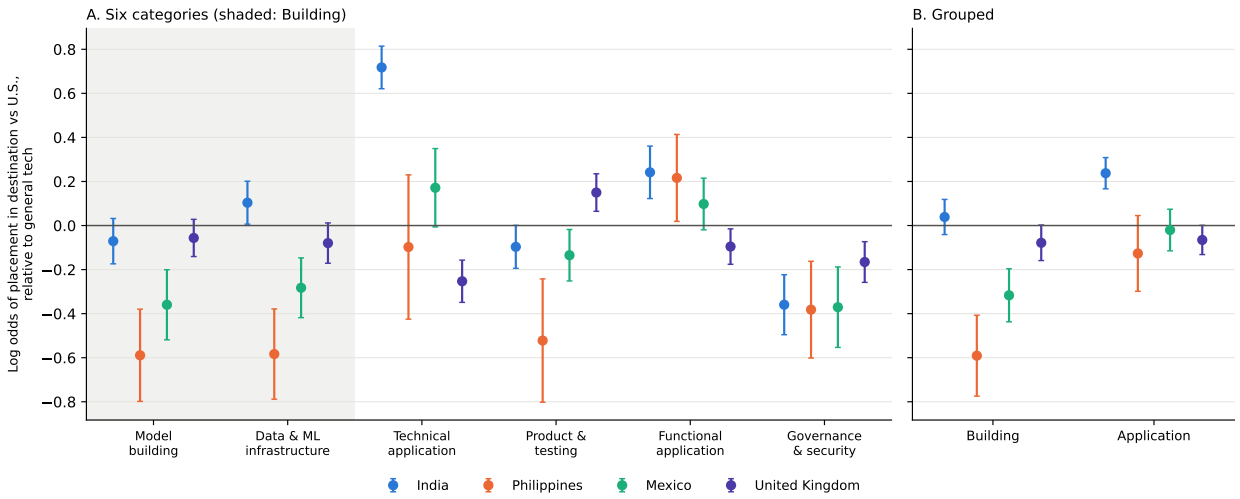



Figure 10: Placement of each task category in the destination rather than the United States, relative to general tech. Panel A: six categories, coefficients $b_{d,k}$ of Equation (2) (shaded: Building). Panel B: Building and Application grouped. All technical jobs, full-text categories, distinct jobs, English-language postings, 2019–2025. 95 percent intervals; standard errors clustered by firm.

skill part is negative where graduates are scarcer than in the United States and zero where supply is the same, so the gap should be negative in the low-wage destinations and can be close to zero in the United Kingdom.

Table 5 reports the destination prices $b_{d,k}$ of Equation (2) for the six focal categories, measured relative to general tech, and Figure 10 plots them. The sample is every technical job of the panel firms in the five countries, 2019–2025: 6.09 million distinct jobs of 4,502 firms, each labelled from its full text (Section 4.5).

Every destination takes some kinds of work more readily than general tech and others less (P1), and each has a mix of its own. India's specialty is technical application. The same firm places a technical-application job in India rather than the United States at odds 2.1 times those of a general-tech job (0.72 log points, s.e. 0.05). India also takes functional application (0.24, s.e. 0.06) and data and ML infrastructure (0.10, s.e. 0.05) more often, governance and security less (−0.36, s.e. 0.07), and model building and product and testing about as often as general tech. The Philippines take functional application more (0.22, s.e. 0.10) and every other category except technical application less, by 0.38 to 0.59 log points. Mexico takes technical application slightly more (0.17, s.e. 0.09) and model building (−0.36), data and ML infrastructure (−0.28) and governance and security (−0.37) less. The United Kingdom takes product and testing more (0.15, s.e. 0.04) and technical application (−0.25), governance (−0.17) and functional application (−0.10) less. All of these are relative odds. In levels the United States remains the larger location for every category, with 3.6 of the 6.1

million distinct jobs.

The comparison the model cares most about is Building against Application (Panel B, Equation (3)). In the three low-wage destinations Building moves out of the United States less: a Building job's odds of being placed in the Philippines rather than the United States are 37 percent lower than an Application job's (−0.46, s.e. 0.05), in Mexico 26 percent lower (−0.30, s.e. 0.04) and in India 18 percent lower (−0.20, s.e. 0.04). In the United Kingdom the gap is small and insignificant (−0.01, s.e. 0.03), and it is the least robust of the four to misclassification (Section 4.4). The gaps arise in different ways. In the Philippines and Mexico, Building itself is placed much less than general tech (−0.59 and −0.32) while Application is placed about like it. India receives Building like general tech (+0.04, s.e. 0.04) and Application more (+0.24), so its gap comes from its strength in Application. By Proposition 3(i), only the skill part of the gap is ordered by graduate supply. The ends of the ranking fit: the Philippines, where doctorates are rarest, has the largest gap, and the United Kingdom has the smallest, which suggests that its location component is also small (Section 6.3 separates the two parts). India fits less neatly: its doctorates are as rare as the Philippines', yet its gap is less than half as large, a point Section 6.4 returns to. Because the United States cancels, the two destinations can also be compared directly. Inside the 694 firms that hire in both in the same year, a Building job's odds of being placed in India rather than the Philippines are 30 percent higher than an Application job's (+0.26, s.e. 0.06); India and Mexico cannot be told apart on this margin (+0.07, s.e. 0.04; Appendix Table 37).

Nor do the destinations share one ranking of the work (P2), as a single index of location costs (Equation (4)) would require. The three low-wage destinations agree in part. Their orderings correlate with India's at 0.43 (the Philippines) and 0.83 (Mexico), and all three put technical or functional application at the top. The United Kingdom orders the categories very differently, with product and testing at the top and technical application at the bottom; its rank correlation is −0.43 with India and −0.66 with the Philippines. A single ranking scaled by distance cannot produce negative correlations. The extended model can, if graduate-intensive work is costlier to place in the low-wage destinations than in the United Kingdom (Equation (15)). Appendix M tests the single ranking formally on measured location-cost requirements and rejects it as well. The single ranking can also be tested as a whole rather than pair by pair. Equation (4) says the $4 \times 6$ matrix of coefficients has rank one, $b_{d,k} = \lambda_d t_k$: one task index, scaled by each destination's distance. Choosing $\lambda$ and $t$ to fit the estimates as closely as their firm-clustered covariance allows and asking how far the fit falls short gives a minimum-distance statistic that is $\chi^2$ with 15 degrees of freedom if the restriction holds. It is 186 ($p < 10^{-30}$; Table 6). The rejection does not rest on the United Kingdom: among India, the Philippines and Mexico alone the statistic is 89 on 10 degrees of freedom, and

Table 6: Do the destinations share one task ranking?

| *A. Minimum-distance tests of one ranking* | $J$ | Degrees of freedom | $p$-value |
|---|---|---|---|
| $b_{d,k} = \lambda_d t_k$ | | | |
| All four destinations | 186.3 | 15 | $< 0.001$ |
| India, Philippines, Mexico | 88.9 | 10 | $< 0.001$ |
| $b_{d,k} = a_d + \lambda_d t_k$ | | | |
| All four destinations | 165.1 | 12 | $< 0.001$ |
| India, Philippines, Mexico | 74.6 | 8 | $< 0.001$ |

| *B. Rank correlation of the six category coefficients* | Spearman | 90% interval | Share of draws $< 0$ |
|---|---|---|---|
| India–Philippines | +0.43 | [+0.14, +0.71] | 0.00 |
| India–Mexico | +0.83 | [+0.49, +0.94] | 0.00 |
| Philippines–Mexico | +0.60 | [+0.43, +0.94] | 0.00 |
| India–United Kingdom | −0.43 | [−0.49, −0.09] | 0.99 |
| Philippines–United Kingdom | −0.66 | [−0.77, +0.03] | 0.95 |
| Mexico–United Kingdom | −0.26 | [−0.54, +0.09] | 0.86 |

*Notes:* The inputs are the 24 coefficients $b_{d,k}$ of Table 5: for destination $d$ and category $k$ = T1–T6, the log odds that a job of category $k$ is placed in the destination rather than the United States, relative to a general-tech job of the same firm and year (PPML, firm × category × year and firm × country × year fixed effects, English postings 2019–2025). One task ranking implies $b_{d,k} = \lambda_d t_k$: a single task index $t_k$ scaled by a destination-specific factor $\lambda_d$. Panel A chooses $\lambda$ and $t$ to minimise $J = (b - \lambda \otimes t)' \hat{V}^{-1} (b - \lambda \otimes t)$, where $\hat{V}$ is the firm-clustered covariance matrix of $b$, with $\lambda$ normalised to one for India; under the restriction $J$ is $\chi^2$ with the stated degrees of freedom. The second restriction adds a destination intercept $a_d$, which allows general tech to sit off the common ranking, and also normalises $t$ to zero for model building. Panel B: Spearman correlation of two destinations' six coefficients at the point estimates; the interval and the share below zero come from 10,000 draws of $b$ from a normal distribution with the estimated mean and firm-clustered covariance. Noise in the draws pulls rank correlations toward zero, so the interval need not be centred on the point value.

allowing each destination its own shift, so that general tech may sit off the common ranking, still rejects in both samples. The largest misfit is the Philippines, whose odds for Building and engineering work fall 0.3 to 0.8 log points short of what the common ranking predicts. The low-wage destinations do broadly agree: their rank correlations are positive in more than 99 percent of draws from the estimates' sampling distribution, while India's correlation with the United Kingdom is negative in 99 percent. The destinations share a rough ordering of the work, then, but not one exact ranking, which is what the extended model, with a location cost and a skill cost that each destination prices differently, allows.

The patterns survive the obvious checks (Appendix O). Measured in percentage points of the probability of placement rather than in log odds, the gap keeps its sign and significance in the three low-wage destinations (Appendix Table 36). Adding the Spanish-language postings moves no gap by more than 0.03 (Table 5, Panel C). Re-estimating on the 510 and 465 firms observed in every year moves no coefficient by more than 1.4 standard errors, and if anything widens the Building gap in the low-wage destinations (India −0.24, Mexico −0.34, the Philippines −0.49) while the United Kingdom's stays at −0.01, so firms entering and leaving the data do not produce it. The choice of titles is the check that moves the estimates. Keeping only the jobs in lookup titles, about a sixth of the technical jobs outside the United States, leaves India (−0.15) and the Philippines (−0.42) close to the main estimates but moves Mexico to −0.11 and the United Kingdom to −0.09. The title-based categories of Appendix H give the same signs in every destination (Table 5, Panel C). Counting the employer's requisitions, which merge a job posted under several records, instead of distinct jobs gives gaps of −0.20 (India), −0.47 (the Philippines), −0.29 (Mexico) and −0.01 (the United Kingdom). And the gaps are not a pattern of Indian IT-services firms. Among the 1,283 firms listed on a U.S. exchange they are −0.21, −0.48, −0.30 and −0.04; firms listed in India show a larger Indian gap (−0.27, s.e. 0.04) and unlisted firms a smaller one (−0.09, s.e. 0.05).

## 6.3 Skill intensity and skill supply (P4)

The prediction tested here is Proposition 2 (Appendix A): the odds of placing a task in a destination fall with its graduate intensity $s_z$ if and only if graduates are scarcer there than in the United States, and the coefficients on $s_z$ are ordered by each destination's relative graduate supply. This is the logic of Costinot (2009) and Costinot and Vogel (2010)—skill-intensive tasks go to skill-abundant countries—applied to jobs inside one firm, and with four destinations the ordering can be checked. Doctoral training is far more common among AI workers in the United Kingdom (18.3 percent hold a PhD) and the United States

(15.5 percent) than in Mexico, and rarer still in India (1.8 percent) and the Philippines (2 percent or less) (Pandit and Nambiar 2026); in 17 of 18 frontier AI firms with enough staff in both countries, Indian technical staff are less likely than U.S. staff to hold a PhD, although the same firms' postings ask for one about as often. The prediction is therefore $\beta^s_{UK} \gtrsim 0 > \beta^s_{MX} > \beta^s_{IN} \gtrsim \beta^s_{PH}$ (P4). The demand shifter works against the prediction: a country with few research employers has a lower premium for given supply, which makes graduate-intensive tasks cheaper to place there, so a negative coefficient for India is evidence that scarce supply dominates, not an artefact of thin demand. Graduate emigration to the United States makes the pattern steeper (Proposition 6 in Appendix B).

Table 7 adds skill intensity to the gravity equation. Skill intensity belongs to the title and the full-text category to the job, so the cells are firm × country × title × category × year, with firm × title × category × year and firm × country × year effects: 1.13 million distinct jobs in lookup titles (the titles for which skill intensity is measured) in 2.14 million cells, of which 1.26 million identify the coefficients. Panel D repeats the main rows with the title-lookup categories for comparison.

The linear version (Panel A) has the sign the model predicts in all four destinations. A title one standard deviation higher in PhD intensity has lower odds of being placed in the destination rather than the United States: the log odds fall by 0.10 in India (s.e. 0.02) and the United Kingdom (0.03), by 0.23 in Mexico (0.04) and by 0.35 in the Philippines (0.06), within 0.02 of the title-lookup estimates. Adding location cost changes neither set of coefficients: India's location-cost slope is −0.26 with or without skill intensity, as expected of two attributes uncorrelated across titles. The gradient also survives destination × full-text category effects (row 3: −0.13, −0.34, −0.26, −0.08), so it comes from titles within the same task category and does not restate the category results of Section 6.2.

That slope averages over a gradient that is far from linear (Panel B). Relative to titles whose U.S. postings never name a PhD, which are mostly business and support roles, titles that name one occasionally (up to 5 percent of postings) have *higher* odds of placement outside the United States in every destination. These are engineering and analytical roles, and the coefficient picks up their technical content. The penalty is at the top. Titles in which more than 30 percent of U.S. postings name a PhD—applied scientists, machine-learning engineers, biostatisticians—have odds of placement outside the United States 30 percent lower in India and in the United Kingdom, 59 percent lower in Mexico and 80 percent lower in the Philippines. On the title-lookup cells the same pattern holds within task categories. Replacing the PhD share by its posting-weighted rank, which gives the many titles with small shares as much weight as the few with large ones, changes the sign (India +0.07, s.e. 0.03; United Kingdom +0.12, s.e. 0.02), for the reason Panel B shows. The data show a

Table 7: Skill intensity and task location, by destination

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| *A. PhD intensity, linear (per standard deviation)* | | | | |
| (1) PhD intensity | −0.099*** | −0.348*** | −0.229*** | −0.104*** |
| | (0.022) | (0.063) | (0.043) | (0.032) |
| (2) With location cost (Appendix M) | −0.119*** | −0.342*** | −0.230*** | −0.103*** |
| | (0.023) | (0.062) | (0.044) | (0.032) |
| Location cost | −0.263*** | −0.149* | −0.032 | 0.027 |
| | (0.050) | (0.083) | (0.067) | (0.048) |
| (3) Within full-text category | −0.129*** | −0.343*** | −0.255*** | −0.083** |
| | (0.021) | (0.060) | (0.048) | (0.033) |
| *B. Share of the title's U.S. postings naming a PhD (reference: none)* | | | | |
| 0–5 percent | 0.258*** | 0.142 | 0.318*** | 0.373*** |
| | (0.079) | (0.108) | (0.100) | (0.051) |
| 5–15 percent | 0.117 | −0.103 | −0.268** | 0.188*** |
| | (0.099) | (0.170) | (0.135) | (0.061) |
| 15–30 percent | −0.015 | −0.410 | −0.380** | −0.101 |
| | (0.146) | (0.261) | (0.183) | (0.088) |
| over 30 percent | −0.349*** | −1.632*** | −0.899*** | −0.362** |
| | (0.123) | (0.322) | (0.248) | (0.162) |
| *C. Building − Application (full-text categories), relative to general tech* | | | | |
| Without PhD intensity | −0.153*** | −0.419*** | −0.111* | −0.093* |
| | (0.057) | (0.085) | (0.067) | (0.049) |
| With PhD intensity | −0.107** | −0.341*** | −0.024 | −0.049 |
| | (0.051) | (0.082) | (0.068) | (0.049) |
| *D. Title-lookup categories (firm × title × year cells)* | | | | |
| PhD intensity, linear | −0.099*** | −0.352*** | −0.241*** | −0.105*** |
| | (0.022) | (0.063) | (0.046) | (0.032) |
| Building − Application, without PhD | −0.095 | −0.526*** | −0.393*** | −0.143* |
| | (0.098) | (0.148) | (0.112) | (0.080) |
| Building − Application, with PhD | −0.042 | −0.434*** | −0.294*** | −0.091 |
| | (0.094) | (0.146) | (0.110) | (0.082) |

*Notes:* PPML on firm × country × title × full-text category × year counts of distinct jobs (English-language postings in lookup titles, 2019–2025), with destination × skill-intensity interactions; fixed effects firm × title × category × year and firm × country × year; standard errors clustered by firm in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. PhD intensity: share of the title's U.S. postings naming a PhD, standardized across titles. Rows (1), (3) and panels B–C: all titles (1,255,883 cells used); row (2): titles outside the lookup's general tech, where location cost is defined (1,185,204 cells used). Row (3) adds destination × full-text category effects. Panel B: one regression; bins fixed before estimation; 46.8, 41.2, 6.3, 2.7, 3.1 percent of U.S. jobs are in the none, 0–5, 5–15, 15–30 and over-30 percent bins. Panel C: destination × Building and destination × Application, without and with destination × PhD intensity. Panel D: the same estimates with title-lookup categories (all distinct jobs, firm × title × year cells).

doctoral margin, a penalty on the most graduate-intensive work, not a gradient that runs through the whole range of tasks.

The penalty at the top survives the checks on the measure in Section 4.6, estimated on the title-lookup cells. Dropping the titles whose PhD share is 0.9 or more, which are the clearest cases of employer templates, gives linear coefficients of −0.07, −0.33, −0.25 and −0.09 for India, the Philippines, Mexico and the United Kingdom. Dropping the top decile of titles by clearance requirements, where substitution ladders are common, gives −0.09, −0.37, −0.23 and −0.08; keeping only titles with at least 100 U.S. postings gives −0.11, −0.48, −0.28 and −0.13. Master's-degree intensity on its own has a weaker gradient of the same sign (−0.05 to −0.32), and when the PhD, master's and research measures enter together, PhD intensity carries the effect.

Proposition 2 says the coefficients should line up with each destination's graduate supply. They are negative in the three low-wage destinations, and the Philippines, where doctorates are rarest, penalizes PhD intensity most, in the linear and in the binned version. The United Kingdom, whose doctoral supply is close to the American, penalizes it less than the Philippines (difference 0.24, s.e. 0.07) and Mexico (0.12, s.e. 0.06), and has the largest coefficient of the four on moderately graduate-intensive work (the 5–15 percent bin).

The ordering does not hold everywhere. The United Kingdom's coefficient is negative, where the model with equal supply predicts zero. One reading is that graduate-intensive work carries a location cost of its own, closeness to U.S. research teams, data and compute, which is common to all destinations and shifts every $\beta_d^s$ down by the same amount. And India's coefficient is as close to zero as the United Kingdom's (difference 0.006, s.e. 0.04), although doctorates are rarer there than in Mexico. India's far larger supply of technical workers with a master's degree may be the margin that matters. I count the ordering test as a partial success.

Model building is the most graduate-intensive category (Section 4.6), so Proposition 3 predicts that part of the Building–Application gap is a skill gap. Panel C of Table 7 asks how much of the gap PhD intensity accounts for. Controlling for destination × PhD intensity reduces the gap from −0.15 to −0.11 in India, from −0.11 to −0.02 in Mexico and from −0.09 to −0.05 in the United Kingdom, by 30, 78 and 47 percent. In the Philippines it falls from −0.42 to −0.34, by 19 percent, just short of the one-fifth threshold fixed in advance. On these titles PhD intensity accounts for almost all of Mexico's small gap and for least of the Philippines'. This is accounting, not a structural decomposition: if PhD-heavy work carries a location cost of its own, as the United Kingdom's coefficient suggests, controlling for PhD intensity removes part of that cost along with the skill premium. Because skill intensity is measured only for lookup titles, these decompositions start from the lookup-title

gaps of Table 5, Panel C; on all technical jobs Mexico's gap is larger, and how much of it is a skill gap is not measured.

The PhD gradient could still be a location cost of research work in disguise rather than a skill premium, and the remaining checks test that reading (Appendix K). Nesting the two is the most direct of them. Starting from destination-specific slopes on the measured location-cost requirements of Appendix M, adding destination × PhD intensity improves the fit (Wald $\chi^2(4) = 94.8$). Starting instead from unrestricted destination × category effects, which absorb any location cost common to a category, it improves the fit as much ($\chi^2(4) = 105.6$) and its coefficients barely move: India from −0.13 to −0.15, the Philippines −0.36 in both, Mexico from −0.24 to −0.27 and the United Kingdom from −0.08 to −0.07. Measured location costs, by contrast, do little for the gap: adding them first moves India's from −0.14 to −0.16 and the Philippines' from −0.40 to −0.41, and they track the unrestricted category effects poorly (correlation 0.34). Research proximity can also be measured directly, as the share of a title's U.S. postings that mention research teams, applied research or publications. Adding it moves the PhD coefficients by at most 0.04 (India from −0.10 to −0.09, the Philippines from −0.35 to −0.33, Mexico from −0.23 to −0.19, the United Kingdom from −0.10 to −0.11), and its own coefficients are jointly insignificant at 5 percent ($p = 0.08$). Differencing against the United Kingdom removes any location cost of graduate work common to all destinations. Net of the United Kingdom, the PhD coefficient is −0.24 (s.e. 0.07) in the Philippines, −0.12 (s.e. 0.06) in Mexico and +0.01 (s.e. 0.04) in India: the supply ordering for the Philippines and Mexico, and not for India. PhD intensity is itself measured from samples of postings, but sampling error does not attenuate it: computing it from two random halves of each title's postings and using each half as an instrument for the other gives the same coefficients to within 0.01.

Section 6.4 follows India's PhD-intensity coefficient year by year.

## 6.4 India's climb up the stack, 2019–2025 (P3, P5)

India is the destination whose placement has changed most, and Proposition 3(iii) says what a climb up the stack should look like. If distance becomes cheaper for all skilled work at once—better tools, larger Indian offices, managers used to working across time zones—India's odds rise for every category above routine support together, and the gap between Building and Application stays where it was. The gap closes only if India's graduate supply rises or the specific cost of placing Building there falls. The data can tell these apart, because the first moves every category against general tech and the second moves Building against Application.

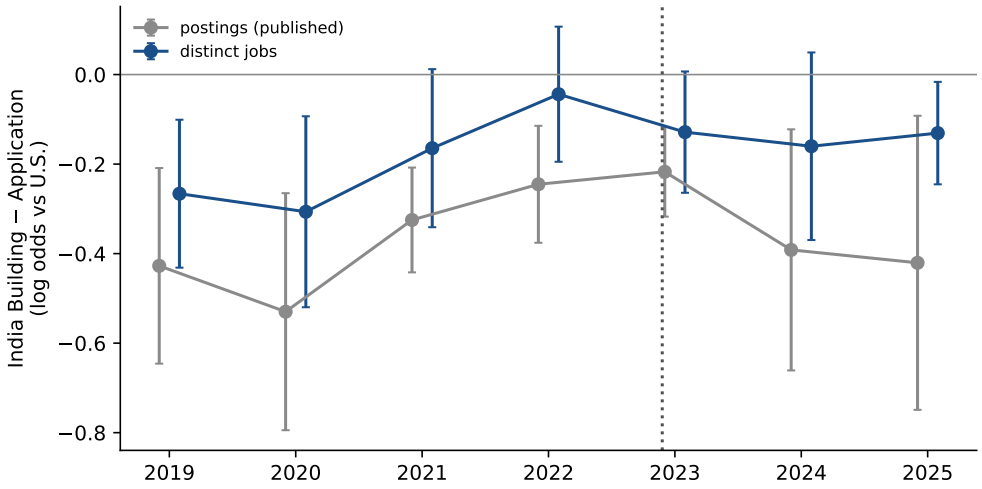



Figure 11: India's Building–Application difference by year, counted in postings and in distinct jobs. PPML coefficients on India × Building minus India × Application, relative to general tech, with firm × category × year and firm × country × year effects; 95 percent intervals, standard errors clustered by firm. Postings: all U.S. and Indian technical postings (the estimates behind Figure 12). Distinct jobs: one per firm × country × title × creation date, among postings with a title in the lookup frame. The dotted line marks the release of ChatGPT.

The counting unit matters here, so I start from distinct jobs, the unit of every other estimate in this section. A U.S. Building job is listed in more cities than a U.S. Application job (2.84 postings against 2.59), while in India the order is reversed (1.66 against 1.98), so counting postings overstates U.S. Building and with it India's shortfall. Figure 11 shows both counts. In distinct jobs among lookup titles, India's Building–Application difference was −0.27 (s.e. 0.08) in 2019, −0.31 (0.11) in 2020, −0.16 in 2021, −0.04 in 2022 and between −0.13 and −0.16 since (−0.13, s.e. 0.06, in 2025); it is significant at the 5 percent level in 2019, 2020 and 2025, at the 10 percent level in 2021 and 2023, and not in 2022 or 2024. On postings it is larger, −0.43 in 2019 and −0.42 in 2025, and significant in every year. Pooled over 2019–2025, the difference is −0.14 (0.06) on jobs counted by day and −0.06 (0.04) on jobs counted by month, against −0.35 (0.08) on postings; on all technical jobs counted by day it is −0.20 (Table 5). Appendix N moves from one estimate to the other one step at a time: the counting unit accounts for almost all of the difference, and the choice between full-text and title labels for at most 0.06. The size of India's gap therefore depends on how jobs are counted, and is between a tenth and a half of the Philippines' however they are counted. Its sign does not: in every year the point estimate says the same firm places an Application job in India more readily than a Building job.

What did change is the level of every category. On distinct jobs, India's Building co-

efficient rose from −0.25 to +0.40 against general tech between 2019 and 2025, and its Application coefficient from +0.02 to +0.53. Figure 12 shows the same climb category by category on postings, the only count available month by month for every category: relative to general tech, India's odds of receiving every other category were higher in 2025 than at the start of the period, with technical application the category India receives most and governance and security the least. The specification, like that of Section 8.1, was written down before it was first estimated on a 2016–2025 window; re-estimated on the paper's 2019–2025 window, every coefficient for 2019–2025 is unchanged. Figure 13 shows the raw counts behind it, which rise and fall with LinkUp's coverage, the 2021–22 hiring boom and the 2023 slowdown, and which is why every estimate compares shares within firm.

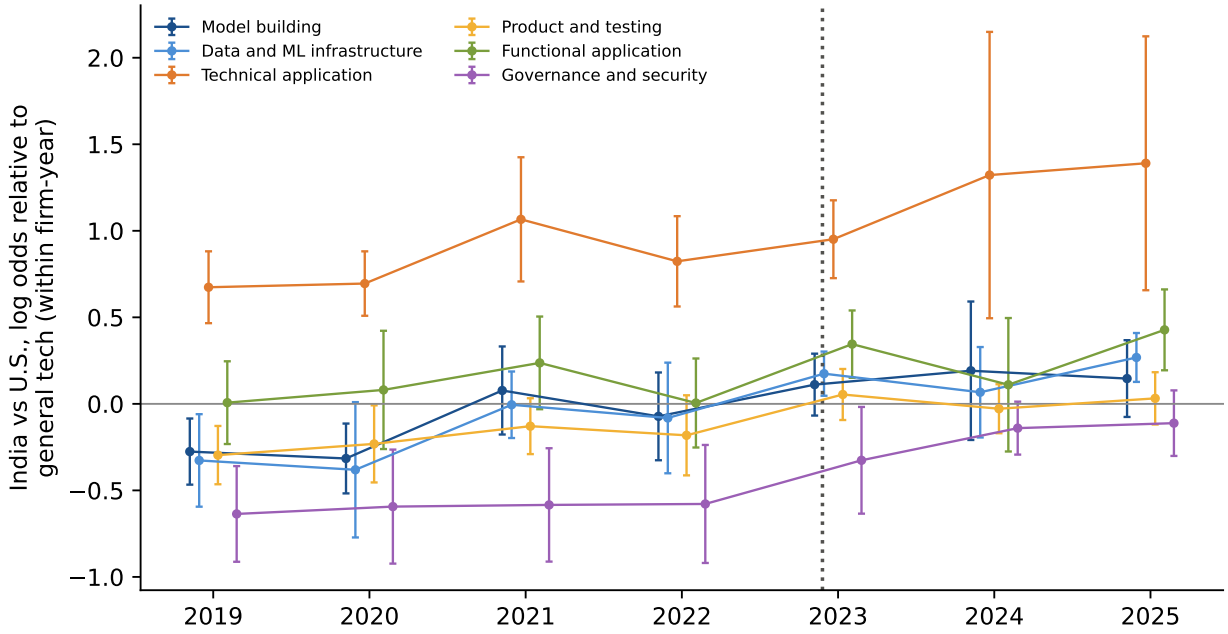



Figure 12: India's odds of receiving each category rather than the United States, relative to general tech, by year, 2019–2025. PPML coefficients on India × category × year with firm × category × year and firm × country × year fixed effects; 95 percent intervals, standard errors clustered by firm. Categories labelled from the full posting text (Section 4.2); 9.1 million U.S. and Indian technical postings, 298,914 cells, 4,494 firms; KPMG, Deloitte and Infosys excluded. The dotted line marks the release of ChatGPT.

Both counts thus show the climb of Proposition 3(iii): Building and Application rose together against general tech, and the gap between them did not close. The firms' Indian offices took more of both without taking over the work the firms keep in the United States. On both counts the gap was smallest in 2022, before ChatGPT, and wider afterwards. The timing fits the shift to remote work in 2020–21, which lowered the cost of directing work from a distance (Dingel and Neiman 2020): relative to general tech, India's odds of receiving both

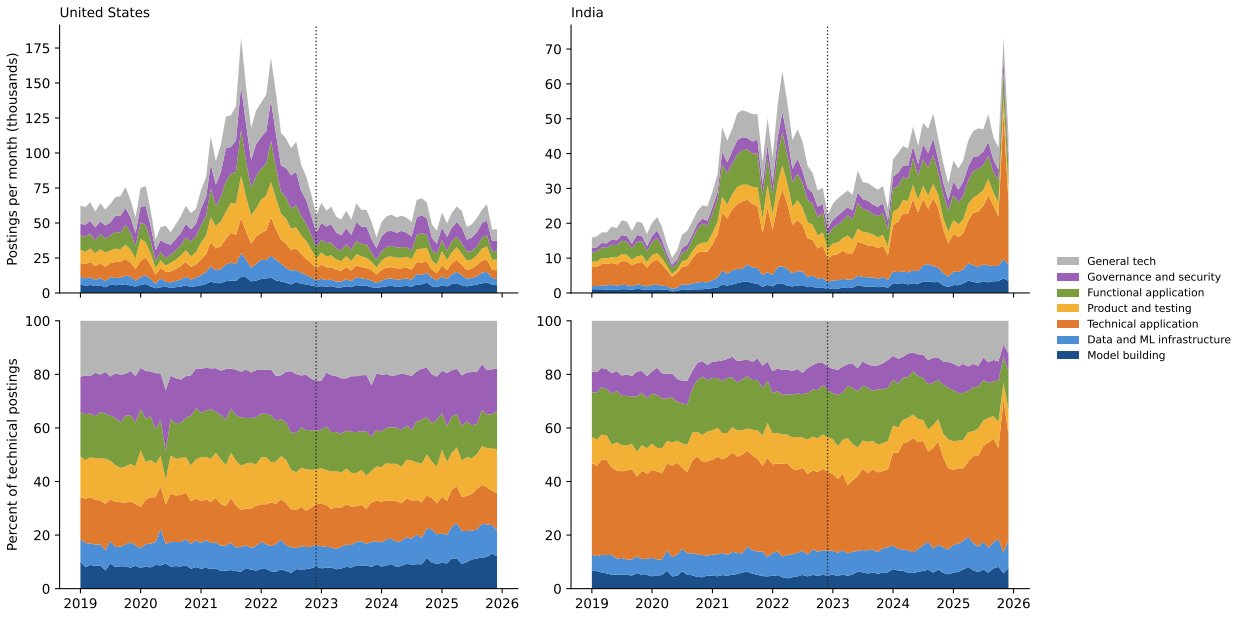



Figure 13: The technology stack in U.S. and Indian postings, 2019–2025. Top: technical postings per month (thousands), stacked by category of the stack. Bottom: each category's share of the country's technical postings that month. Categories are read from the full posting text (Section 4.2); Building (model building, data and ML infrastructure) is at the bottom of each stack. 9.1 million postings of the panel firms; KPMG, Deloitte and Infosys excluded. Counts mix hiring with LinkUp's scraping coverage, which grew over the period, especially in India; the last month (December 2025) shows a jump in Indian postings that may reflect a collection vintage. The dotted line marks the release of ChatGPT.

Building and Application work jumped in those years (Section 8.1). In the model a fall in the cost of distance common to all skilled work raises every category together and leaves the Building gap open, which is what the data show. The comparison has no untreated group, so it describes timing rather than cause, and part of the rise reflects LinkUp's expanding coverage of India.

The skill side moved in the same years. Estimated year by year on the title-lookup cells, India's PhD-intensity coefficient rose from −0.24 (s.e. 0.06) in 2019 to −0.17 in 2020, −0.12 in 2021 and −0.06 in 2022, and has stayed between −0.07 and −0.12 since (−0.07, s.e. 0.02, in 2025). A general fall in the cost of distance cannot produce this, because in Equation (9) it is absorbed by $\gamma_{ict}$ and leaves $\beta^s_{dt}$ unchanged. A rise in India's relative supply of graduates able to do the work can, and so can a fall in the location cost specific to graduate-intensive work; these data do not separate the two. The flattening came early and stopped, which fits the Building gap's own path. The supply side may help explain why the gap persists. The companion paper finds that 62.5 percent of India's AI workers stopped at a bachelor's degree, and that 56 percent of India-trained AI workers with a master's degree live in the United States, against 43 percent of other India-trained technical workers (Pandit and Nambiar 2026). In the model, that emigration holds India's graduate supply down from both sides (Proposition 6).

## 6.5 Inside the firm

A last comparison asks when in a firm's history Building work is placed in India, following firms from the year they first appear there. If a firm's Indian office starts with Application and adds Building as it learns, a firm-specific cost of placing Building there falls with tenure, and the firm's own gap should narrow. I date a firm's entry by the first year with at least five Indian technical distinct jobs and estimate the gap separately by years since entry for the 1,338 firms that enter in 2020–2024, with the 757 firms already present in 2019 as a separate group (Table 8 and Figure 14). The data do not show such a sequence. Entrants' gap is −0.10 (s.e. 0.06) in the entry year and −0.05 (0.07) three or more years later, a difference of +0.05 (0.09); it rises in the second year and falls back. Dating entry at one job or at twenty, dropping the entrants that had posted Indian jobs of any occupation before their technical entry, or keeping only firms with many U.S. jobs gives the same answer. If anything, the firms observed longest in India have the widest gap (−0.23, s.e. 0.04), which is the opposite of a climb inside the firm. Entry into the data partly reflects LinkUp's growing coverage of Indian career sites, and nearly half the entrants had posted Indian jobs of other kinds earlier, so this evidence is descriptive and cannot rule out learning that happened before a

Table 8: Building − Application gap in India by years since the firm's first year with Indian technical postings

| | Years since entry | | | | Present | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3+ | from 2019 | 3+ − 0 | Entrants |
| Main: entry at 5 jobs | −0.10 | −0.01 | +0.14** | −0.05 | −0.23*** | +0.05 | 1,338 |
| | (0.06) | (0.06) | (0.07) | (0.07) | (0.04) | (0.09) | |
| Entry at 1 job | −0.13* | −0.03 | +0.15** | −0.05 | −0.23*** | +0.08 | 1,751 |
| | (0.07) | (0.07) | (0.07) | (0.07) | (0.04) | (0.10) | |
| Entry at 20 jobs | −0.10* | −0.07 | +0.07 | −0.07 | −0.24*** | +0.03 | 925 |
| | (0.05) | (0.06) | (0.07) | (0.07) | (0.05) | (0.08) | |
| Earlier Indian postings dropped | −0.19*** | −0.10 | +0.02 | −0.19*** | −0.23*** | +0.00 | 703 |
| | (0.07) | (0.06) | (0.08) | (0.06) | (0.04) | (0.09) | |
| 50+ U.S. jobs | −0.11 | −0.01 | +0.14** | −0.05 | −0.24*** | +0.06 | 891 |
| | (0.07) | (0.06) | (0.07) | (0.07) | (0.04) | (0.09) | |

*Notes:* Each row is one Poisson pseudo-maximum-likelihood regression on cells of firm × country (United States, India) × group (Building, Application, general tech) × year, 2019–2025, English distinct jobs, with firm × group × year and firm × country × year fixed effects. Entries are the Building coefficient minus the Application coefficient: the log odds of placing a Building job in India rather than the United States minus the same log odds for an Application job, inside the same firm and year (general tech is the common base and cancels). A firm enters in the first year with at least five Indian technical distinct jobs (one or twenty in the second and third rows). Firms already over the threshold in 2019 are pooled in the column *Present from 2019*; entrants are firms entering in 2020–2024 with U.S. postings in the entry year; their Indian cells before entry are left out. *Earlier Indian postings dropped*: entrants with Indian postings of any occupation before their entry year are excluded. *50+ U.S. jobs*: firms with at least 50 U.S. technical distinct jobs over 2019–2025. Entry is entry into the data, which also reflects the expanding coverage of Indian career sites, so the table is descriptive. Standard errors clustered by firm in parentheses, from the joint covariance matrix. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

firm appears. It does fit the account of Section 6.4: India's climb is common to the firms, not a sequence each firm repeats, and what holds Building in the United States does not wear off with a firm's time in India.

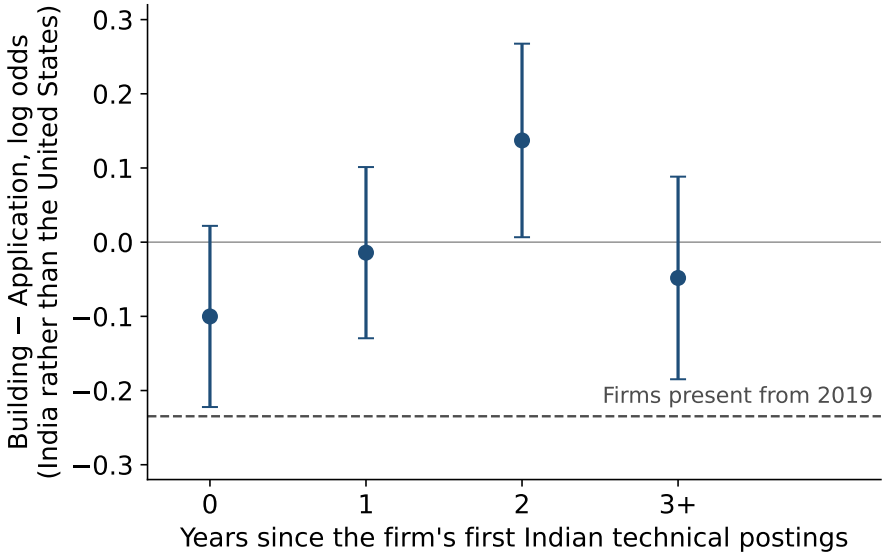



Figure 14: India's Building–Application gap by years since the firm's first year with at least five Indian technical distinct jobs. PPML coefficients on India × Building minus India × Application by years since entry, relative to general tech, with firm × group × year and firm × country × year effects; 95 percent intervals, standard errors clustered by firm. Entrants: 1,338 firms entering in 2020–2024; dashed line: the 757 firms present from 2019. English distinct jobs, U.S. and India, 2019–2025; KPMG, Deloitte and Infosys excluded. Entry into the data also reflects the expanding coverage of Indian career sites.

# 7 Did U.S. pay fall behind?

Firms place the codified, application side of their technical work in India more readily than model building. The obvious worry is the U.S. workers who still do that application work. In GRH the cost saving from moving tasks goes to them (the productivity effect of Section 5.1), while workers whose tasks moved bid down pay in the tasks that remain (the labor-supply effect). The model's test is one-sided (P7): if the productivity effect dominates, pay in the work that moved does not fall behind comparable work. This section reports that test. It is descriptive and short, because nothing here separates the movement of tasks from the other forces that moved technical pay in these years. By Proposition 5, the fall in the U.S. share of the firms' jobs between 2019 and 2025 would be worth 9 to 19 log points of pay for Application and 10 to 22 for Building at fixed prices, for placement elasticities $\theta$ between

2.2 and 4.8; these data do not identify $\theta$, and part of the fall is LinkUp's growing coverage of India, so the numbers are an illustration (Appendix Q).

Pay can be measured as the wage actually paid or as the salary advertised. The Occupational Employment and Wage Statistics (OEWS) of the Bureau of Labor Statistics report wages actually paid, by occupation; I weight each occupation by its share of each group's U.S. jobs in LinkUp, so this measure cannot tell Building and Application apart where they share occupations. WageScape records the salary advertised in U.S. postings and follows the same title over time, so a change in its pay is not a change in the mix of jobs.

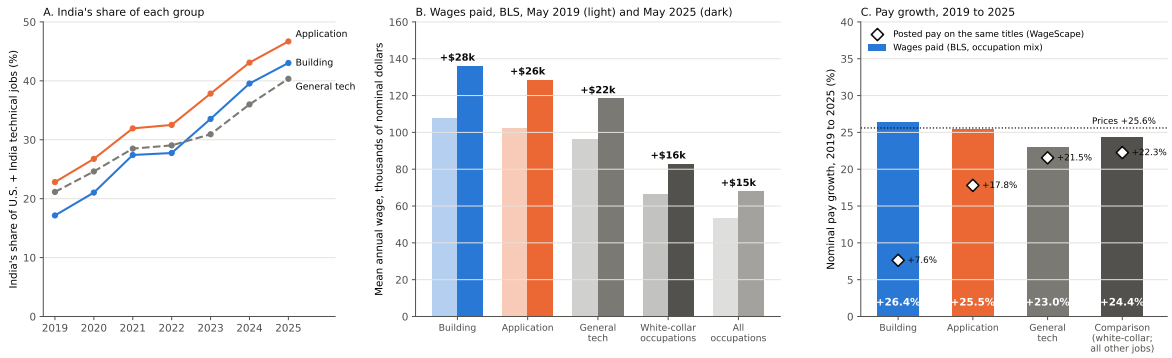



Figure 15: India's share of the work and U.S. pay, 2019–2025. Panel A: India's share of the firms' U.S.-plus-India technical jobs in each group (distinct jobs, full-text categories; part of the rise reflects LinkUp's growing coverage of India). Panel B: BLS OEWS mean annual wage, May 2019 and May 2025; Building, Application and general tech weight occupations by the O*NET mix of each group's U.S. jobs; white-collar is SOC major groups 11–29, 41 and 43, employment-weighted; all occupations is the OEWS national mean. Panel C: the same growth in percent (bars), and WageScape posted pay on the same titles, 2019 to 2025 (diamonds), whose comparison is all other U.S. jobs. Dotted line: consumer prices (CPI-U), May 2019 to May 2025. KPMG, Deloitte and Infosys excluded.

Wages actually paid kept pace. While India's share of the firms' advertised technical jobs roughly doubled (Figure 15, Panel A), the mean annual wage in the occupation mix of U.S. Building jobs rose from $107,800 in May 2019 to $136,200 in May 2025, and in Application from $102,500 to $128,700. These gains of 26.4 and 25.5 percent match the 24.4 percent of white-collar occupations as a whole and the 25.6 percent rise in consumer prices. Pay in the occupations of that work did not fall behind comparable work, and its real level did not fall.

Advertised pay on the same titles grew less. Relative to all other U.S. jobs, posted salaries fell behind by 4.4 log points in Application titles (s.e. 1.5) and by 13.7 in Building titles (s.e. 2.9), and on the titles observed in both 2019–20 and 2024–25 they fell in real terms (Figure 16). Most of this shortfall reflects where pay started. Building titles start far higher ($103,800 against $44,300 for other jobs), and across the whole U.S. market better-paid titles got smaller raises in these years, as tight post-pandemic labor markets lifted pay

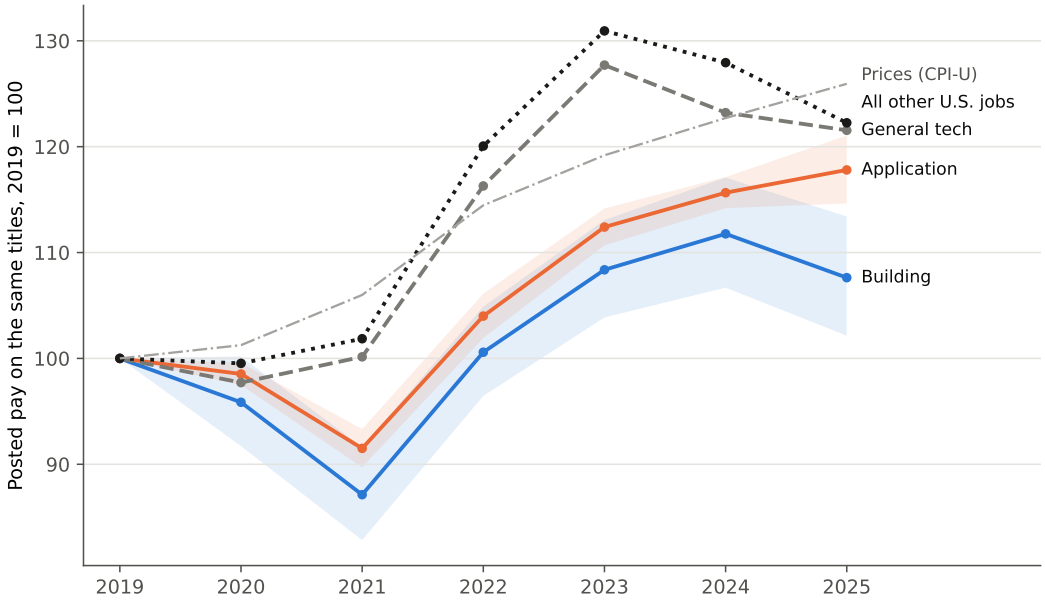



Figure 16: Posted pay on the same job titles, by year, 2019 = 100. WageScape salaries regressed on year indicators with title fixed effects, separately for each group; 95 percent bands, clustered by title, for Building and Application. Titles are assigned to a group by the full-text category of most of their U.S. LinkUp jobs. "All other U.S. jobs" is a 2 percent random sample of other WageScape postings, weighted up. Prices: CPI-U, annual average (2025: January–July).

at the bottom fastest (Autor et al. 2023); Building's posted pay ran below other jobs from 2020, before most of the rise in India's share. The part linked to India is small. Across 1,552 titles, posted pay grew less where India's share rose more, by 8.8 log points per unit of India share once starting pay is held fixed (s.e. 3.4): a title whose Indian share rose by ten points saw pay grow about 0.9 percent less, a twentieth of the 18 percent rise in posted Application pay. That is the sign of a labor-supply effect, and these data do not rule one out.

The test therefore passes in its one-sided form: wages paid in the occupations of the work that moved kept pace with other white-collar pay. The evidence cannot show a net gain relative to other workers, because under one white-collar labor market the gain is shared with workers whose tasks never moved.

# 8 Generative AI as a test of the model

ChatGPT, built on OpenAI's GPT-3.5 model, was released in November 2022, when these firms' division of technical work across countries was already in place. The model makes two predictions about such a technology, and they concern different margins of the same data (Proposition 4). If language models make distant work easier to direct and check, they lower the cost of distance and should move work out of the United States, starting with the codified tasks already close to the margin. If they mostly do codified work themselves wherever it is done, they should leave the division of work where it was and change how much of each task firms advertise, starting with juniors. This section looks at both margins in the same firms' U.S. and Indian offices.

## 8.1 Did generative AI move work out of the United States? (P6)

If the release made it cheaper to do technical work in India, India's share of each kind of work should jump at the end of 2022. Shares move for many reasons, so a jump means something only if it is larger than the jumps the same data show at dates when nothing special happened. I estimate the jump in India's odds of receiving each category, relative to general tech and beyond the category's own trend, at the fourth quarter of 2022 and at seven earlier quarters, and ask where the real date ranks among the eight. The data are counts of U.S. and Indian postings by firm, country, category and quarter from 2019 to 2025, with every comparison inside the same firm and quarter. The design and the prediction of no significant jump were written down before estimation, in a pre-analysis plan dated 15 October 2026 that used a window starting in 2018; the estimates here use the paper's 2019–2025 window, with the placebo dates moved a year later so that each keeps four quarters on

either side. The details are in the table notes.

Table 9: Estimated changes in India's relative task allocation around ChatGPT's release

| Category | Break at 2022Q4 | Conv. $p$ | Mean \|placebo\| | Rank $p$ |
|---|---|---|---|---|
| *Building* | | | | |
| T1 Model building | 0.070 | 0.59 | 0.27 | 0.88 |
| | (0.130) | | | |
| T2 Data and ML infrastructure | 0.104 | 0.58 | 0.20 | 0.88 |
| | (0.189) | | | |
| *Application* | | | | |
| T3 Technical application | −0.007 | 0.96 | 0.35 | 1.00 |
| | (0.132) | | | |
| T4 Product and testing | 0.165 | 0.23 | 0.12 | 0.50 |
| | (0.138) | | | |
| T5 Functional application | 0.135 | 0.40 | 0.23 | 0.75 |
| | (0.161) | | | |
| T6 Governance and security | 0.175 | 0.22 | 0.05 | 0.13 |
| | (0.142) | | | |

*Notes:* PPML on firm × country × category × quarter counts of U.S. and Indian technical postings, 2019Q1–2025Q4, categories from the full posting text; fixed effects firm × category × quarter and firm × country × quarter; regressors India × category, India × category × trend and India × category × Post (2022Q4 on), reference category general tech. Standard errors clustered by firm in parentheses. Placebo breaks at each quarter 2020Q1–2021Q3, each estimated on 2019Q1–2022Q3 (7 dates); rank $p$: share of the 8 dates with an absolute break at least as large as the true one (smallest attainable value 0.125). Pre-specified and first estimated on 2018Q1–2025Q4 (11 placebo dates, breaks 0.10 to 0.27, rank $p$ 0.08 to 0.92); re-estimated on the paper's 2019–2025 window after those results were seen; both are reported. Stars use the conventional $p$: $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

No category's jump at the end of 2022 is significant (Table 9; conventional $p$ from 0.22 to 0.96), and for five of the six it lies within the range of the placebo jumps. For technical application, the work that code-writing tools reached first, it is −0.01 (s.e. 0.13), and every one of the seven placebo dates shows a jump at least as large. Governance and security is the exception, and its rise shows up throughout the data. Its jump of 0.175 (s.e. 0.142) is larger than all seven placebo jumps (rank $p$ 0.125, the smallest value seven dates allow), and India's odds of receiving it, relative to general tech, rose from −0.58 in 2022 to −0.33 in 2023 and −0.11 in 2025 (Figure 12). It is plausibly work that generative AI created rather than automated. Reviewing a model for risk before it is hosted (Section 4.2) concerns models, but its duties are auditing, not building, so it can be done away from the teams that train them. The jump is not significant by its own standard error, and I read it as the start of a shift rather than a break. With seven placebo dates the rank test cannot reject at the 10 percent level, so the conclusion for the other categories rests on the conventional standard errors.

No significant break is not the same as no change. The estimates are imprecise: for technical application the 95 percent interval runs from −0.27 to +0.25, so India's odds of receiving it, relative to general tech, could have risen by as much as 29 percent or fallen by 23 percent, and for the other categories the upper ends of the intervals are rises of 38 to 61 percent. What the table shows is that the design detects no category moving toward India faster than general tech around the release.

The same firms' non-technical postings measure part of what that comparison cannot (Appendix Table 31). Non-technical work shares the coverage of the firm's Indian career site and the firm's other shocks, but not changes in the cost of doing technical work far away, so comparing all technical jobs with it picks up an India-ward shift common to every technical category. At the release that shift is +0.08 (s.e. 0.14), smaller than at five of the seven placebo dates (rank $p$ 0.75); against white-collar non-technical postings alone it is 0.00 (s.e. 0.11). The intervals still allow a rise in India's odds for technical work of up to 41 percent (25 percent against white-collar work), so a moderate common shift cannot be ruled out, but none is visible. Changes of unit and sample give tighter answers for the six categories against general tech. On monthly counts with the break in November 2022, the pooled break is 0.11 (s.e. 0.12), at rank $p$ 0.82 among 21 monthly placebos. Counting distinct jobs instead of postings it is 0.01 (s.e. 0.06), and counting requisitions 0.01 (s.e. 0.06). Among the 510 firms observed in every year, whose coverage changed least, it is 0.00 (s.e. 0.06), which rules out a differential rise in India's odds of more than 11 percent.

The test is not blind to real shifts. The largest jumps in the same series come in 2020 and 2021, when work went remote: up to +0.7 log points, with +0.70 (s.e. 0.24) for technical application and +0.45 (s.e. 0.18) for model building at the end of 2020. That is what a fall in the cost of doing work away from the office looks like in these data when it is not the same for every task. The comparison was not planned in advance and needs its own design, but it shows that differential jumps of this size can be seen. No jump of that size is detected after ChatGPT.

Why might the design detect no shift after generative AI arrived? In the model a job moves only when its cost falls in one country relative to another. The tools that spread in 2023–25 mostly do part of the work themselves—they write code, draft documents and summarize meetings—and save those hours wherever the job is done, so, like the coding assistant that left the bank's phone app in Bengaluru at $31.50 against $70 (Proposition 4(ii)), they leave India exactly as much cheaper as before. The large fall in the cost of distance itself had come earlier, with remote work, and by the end of 2022 much of the work that is cheap to direct from afar had already been placed. In the model, what is advertised mainly in the United States is held there by doctoral skill and by location costs specific to that work,

and a tool that saves labor in every location changes neither. The design detects only shifts that differ across categories, however, and the period since the release is short, so a slower relocation, as firms learn to use the tools, would not yet be visible in these data.

## 8.2 Which software-developer postings fell after ChatGPT

The evidence on generative AI and U.S. hiring centres on software developers, and the studies of it (Westby et al. 2026; Brynjolfsson et al. 2025a; Hosseini and Lichtinger 2025) treat them as one occupation. Exposure is likewise measured from the tasks of occupations, as in the vacancy study of Acemoglu et al. (2022), or from the tasks a posting advertises, as in Shukla (2026); the full-text categories read the task from each job instead. The occupation is not one kind of work, however: in the United States about half of its jobs are technical application, the coding of products and systems, and 17 percent are Building, the design and training of models and the data systems behind them (Figure 8, Panel B). Code assistants write code; they do not decide what a model should learn or how its data should flow. If they are behind the change in hiring, the decline should be concentrated among the software developers who code applications, and those who build models and data systems should be spared. The full-text categories make it possible to ask which software-developer postings fell.

I compare each kind of software-developer job with the same firm's general-tech jobs in the same quarter. For firm $i$, group $g$ and quarter $q$ (three-month blocks counted from November 2022), $V_{igq}$ counts distinct jobs in three groups: software developers (O*NET 15-1252) doing Application work (T3–T6), software developers doing Building work (T1–T2), and general-tech jobs in any technical occupation (T7). I estimate

$$\mathbb{E}[V_{igq}] = \exp\Big( \sum_{h \neq \mathrm{GT}} \sum_{k \neq -1} \beta_{h,k} \mathbb{1}[g = h] \mathbb{1}[q = k] + \alpha_{iq} + \gamma_{ig} \Big), \tag{6}$$

by PPML with general tech as the omitted group. The firm × quarter effect $\alpha_{iq}$ takes out everything that moves all of a firm's hiring in a quarter, including the 2021–22 boom, the 2023 slowdown and LinkUp's coverage of the firm. The firm × group effect $\gamma_{ig}$ takes out each firm's usual mix of the three groups. What is left, $\beta_{h,k}$ for group $h$, is how the ratio of a firm's software-developer jobs of kind $h$ to its general-tech jobs moved in quarter $k$, relative to August–October 2022. General tech is the comparison because it is the work that code assistants reach least and because its demand follows the firm's overall technical activity.

Figure 17 shows the answer. In the United States, Application jobs of software developers move in step with general tech for the fourteen quarters before the release: the pre-trend is flat (joint test of the leads, $p = 0.22$). They then fall by about 0.07 log points in the first

Table 10: Software-developer jobs against general tech after ChatGPT, within firm

| | Pooled | De-trended | Placebo (Nov 2021) | Leads $p$ |
|---|---|---|---|---|
| *A. United States* | | | | |
| Software developers, Application (T3–T6) | $-0.157^{***}$ | $-0.146^{***}$ | 0.007 | 0.22 |
| | (0.030) | (0.042) | (0.037) | |
| of which technical application (T3) | $-0.158^{***}$ | $-0.108^{**}$ | $-0.027$ | 0.07 |
| | (0.035) | (0.043) | (0.048) | |
| Software developers, Building (T1–T2) | $-0.078$ | $-0.063$ | $-0.043$ | <0.001 |
| | (0.076) | (0.053) | (0.048) | |
| Application in other technical occupations | $-0.113^{***}$ | $-0.071^{**}$ | | |
| | (0.028) | (0.029) | | |
| Software developers minus other occupations | $-0.032$ | $-0.070^{**}$ | | |
| | (0.025) | (0.030) | | |
| *B. India* | | | | |
| Software developers, Application (T3–T6) | $-0.027$ | $-0.110^{**}$ | 0.020 | 0.32 |
| | (0.050) | (0.047) | (0.056) | |
| of which technical application (T3) | $-0.013$ | $-0.113^{*}$ | $-0.003$ | 0.02 |
| | (0.061) | (0.065) | (0.087) | |
| Software developers, Building (T1–T2) | 0.092 | $-0.142^{***}$ | $-0.025$ | 0.12 |
| | (0.066) | (0.050) | (0.056) | |
| Application in other technical occupations | 0.036 | $-0.034$ | | |
| | (0.057) | (0.059) | | |
| Software developers minus other occupations | $-0.081$ | $-0.107^{**}$ | | |
| | (0.054) | (0.043) | | |

*Notes:* PPML on firm × group × quarter counts of distinct jobs, English-language postings, February 2019 – October 2025, with firm × quarter and firm × group fixed effects; the omitted group is general tech (T7, any technical occupation). Coefficients are log points relative to general tech after November 2022. "De-trended" adds a group-specific linear quarter trend. "Placebo" moves the release to November 2021 and uses only quarters before August 2022. "Leads $p$": joint test of the 14 pre-release quarters of the event study (Figure 17). The last two rows of each panel come from one regression that adds Application and Building jobs in other technical occupations as separate groups. Software developers: O*NET 15-1252. Standard errors clustered by firm in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$. This specification was chosen after seven pre-specified comparisons with other technical occupations showed only a pre-existing trend; see Section 8.2.

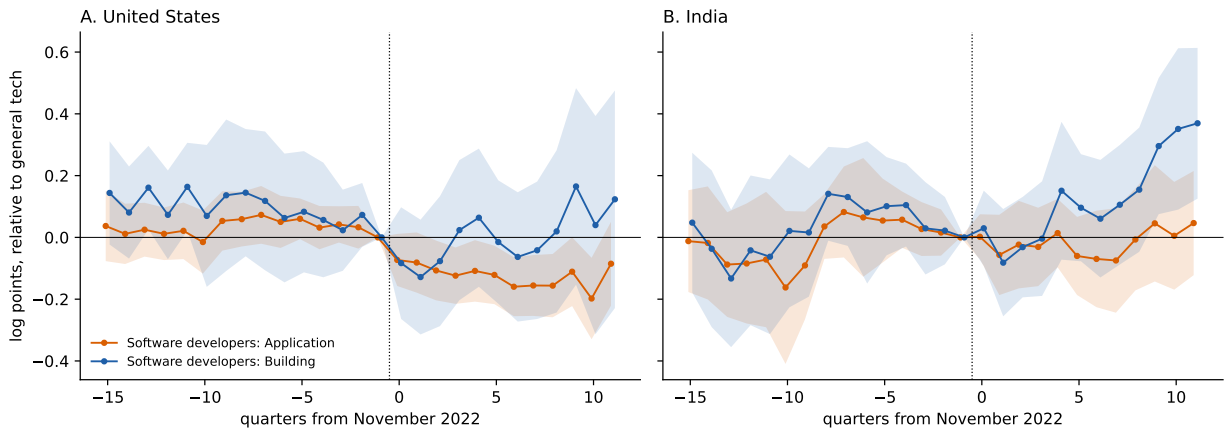



Figure 17: Software-developer Application and Building jobs relative to the same firm's general-tech jobs, by quarter from November 2022. Coefficients $\beta_{g,k}$ of Equation (6), PPML with firm × quarter and firm × group effects, reference August–October 2022 (set to zero), 95 percent intervals, standard errors clustered by firm. Software developers: O*NET 15-1252; Application: T3–T6; Building: T1–T2; general tech: T7 in any technical occupation. Distinct jobs, English-language postings, February 2019 – October 2025.

quarter after the release, and keep falling to about −0.15 over the next two years, where they stay through 2025. Pooled, the decline is 0.157 log points (s.e. 0.030), or 15 percent relative to general tech (Table 10). It survives a group-specific linear trend (−0.146, s.e. 0.042), and moving the release to November 2021 gives nothing (+0.007). The coding slice alone, technical application, falls by as much (−0.158, s.e. 0.035), although its leads are borderline ($p = 0.07$), short of the pre-specified bar of 0.10. Building jobs of software developers show no lasting decline. They dip by 0.08 to 0.13 log points in the three quarters after the release, about as much as Application at first, but the dip is imprecise and gone by the end of 2023; pooled, the change is −0.078 (s.e. 0.076), and −0.063 (s.e. 0.053) once de-trended. The comparison is weaker for Building than for Application. Its ratio to general tech had been falling for three years before the release, and the leads are jointly significant ($p < 0.001$), so the Building estimates describe what happened rather than test a break.

This complements Westby et al. (2026), who study a different margin of the same change. Using the near-universe of U.S. online vacancies, they show that junior software-developer postings fell by 16.3 percent relative to senior ones in the year after ChatGPT (their difference-in-differences; their triple difference against other occupations, the comparison used in Section 8.3, is −0.136), and they attribute the fall to the routine coding that junior developers do and code assistants perform, while senior developers do more creative and complex work. The estimates here find a decline of similar size in application work, most of it coding, inside the same firm and quarter, and no lasting decline in the design

and training of models and data systems. The two results are different margins of the same change: the task measure shows which work fell, and their experience comparison shows a separate, shorter tilt against junior postings.

The decline follows the task, not the job title. Application jobs in other technical occupations fell against general tech as well ($-0.113$, s.e. 0.028), by nearly as much as those of software developers; the gap between the two is $-0.03$ (s.e. 0.025) pooled and $-0.07$ (s.e. 0.030) de-trended, while Building jobs in other occupations did not fall ($+0.003$). What fell after ChatGPT was application postings wherever they sit, and software developers' postings fell most because most of their work is application. This is the paper's measurement point applied to generative AI: an occupation-level comparison sees a decline among software developers, while the task-level measure shows which of them, and that the same work fell in other occupations too. And India shows no comparable pooled decline: the change for Indian software-developer Application jobs is $-0.027$ (s.e. 0.050). Once group-specific trends are removed, Indian Application and Building jobs of software developers both fall ($-0.110$ and $-0.142$), so in India any change is not specific to application work. That is consistent with Section 8.1, which detects no differential shift toward India: the change detectable after the release is in what is advertised in the United States. The comparison with general tech was chosen after a pre-specified comparison of software developers with all other technical occupations. Because application work fell in those occupations too, that comparison shows only the slow decline of the coding share among U.S. software developers that had begun before 2020 (from 50.4 percent in early 2019 to 46.0 percent in late 2022; leads $p = 0.021$, placebo in November 2021 $-1.92$ points against $-1.46$ at the release).

The experience ladder gives a second view of the same firms. Friebel et al. (2026) model a firm's internal labor market as a pyramid: juniors perform routine tasks and learn on the job, and the best of them become the seniors who run the firm's complex work. In their model AI can act in two ways. If it raises productivity, firms freeze junior hiring for a time, the pyramid bulges toward a diamond, and in the long run the pyramid returns. If it speeds up learning, so that fewer seniors are needed to handle what mid-level workers can now solve, the pyramid turns permanently into a diamond. Which work fell is one question; whether firms also changed how junior the remaining work is allowed to be is another, and the natural next one: whether firms cut junior software hiring, whether the change was temporary or lasting, and whether the ladder changed shape in the same way in the United States and in India.

Table 11: What is different about junior software postings: the mix of work or the requirement for the same work?

| | | By task category | | By firm |
|---|---|---|---|---|
| | Difference | Mix of work | Same work | Mix of firms |
| *A. India minus the United States* | | | | |
| Before ChatGPT (Nov 2021–Oct 2022) | −11.9 | 0.7 | −13.1 | −5.8 |
| First year after (Nov 2022–Oct 2023) | −13.5 | −0.6 | −13.3 | −3.3 |
| 2024–2025 | −19.3 | 0.2 | −18.0 | −4.5 |
| *B. Change from before ChatGPT, United States* | | | | |
| To the first year after | −6.4 | 0.3 | −6.8 | −4.6 |
| To 2024–2025 | −2.8 | 0.4 | −3.0 | 1.3 |
| *C. Change from before ChatGPT, India* | | | | |
| To the first year after | −8.1 | −0.6 | −7.8 | −6.2 |
| To 2024–2025 | −10.1 | 0.8 | −9.9 | −6.7 |
| *D. Same firm and same title in both countries: before / first year after / 2024–25* | | | | |
| Junior share, India minus U.S. | | −14.8 / −10.9 / −11.8 | | |
| Years of experience asked, India minus U.S. | | 1.12 / 0.99 / 0.87 | | |

*Notes:* Software-developer postings (O*NET 15-1252) of the panel firms with a stated minimum experience requirement; junior = 0–2 years; percentage points. "Difference" is the change or gap in the junior share over all postings. Shift-share as in Westby et al. (2026): the junior share is a weighted average over job types, and its change splits into a part from posting a different mix of types ("mix") and a part from a different junior share within the same type ("same work"); the small interaction term is not shown. Task categories are the full-text categories, which cover every posting. The firm decomposition uses firms with at least 20 postings in both groups (50 to 95 percent of postings). Panel D: firm × title pairs with at least five postings in each country in the period, weighted by U.S. postings (302 to 753 pairs). KPMG, Deloitte and Infosys excluded. Descriptive.

### 8.3 Junior software postings

Appendix D applies the design of Westby et al. (2026) to the firms' U.S. and Indian offices separately, comparing junior with senior software-developer postings before and after November 2022 and netting out the same comparison in the firms' other technical occupations. In the U.S. offices junior software postings fell by 19 percent (−0.215, s.e. 0.063); with Westby et al.'s own two-rung split the estimate is −0.155 (s.e. 0.052), close to their −0.136. The decline was short-lived, concentrated in 2023 and gone by 2024: the temporary freeze in junior hiring that Friebel et al. (2026) predict when AI raises productivity. The Indian offices show no such decline (−0.021, s.e. 0.073). Nor does India's lower junior share reflect a different mix of tasks; a shift-share decomposition shows that the same title at the same firm simply asks for about a year more experience in India (Table 11). After ChatGPT the junior share fell within tasks in both countries, but by 2024–25 it had recovered in the United States and not in India, where the software ladder had shifted toward mid-level jobs, a shift that began before the release.

This junior dip should not be confused with the fall in application work documented in Section 8.2; the two are separate adjustments. The fall in application work runs through all experience levels and has persisted, whereas the junior dip was brief and, split by task, lies in software developers' Building jobs (−0.41, s.e. 0.12) rather than their Application jobs (−0.10, s.e. 0.06), although the same Building comparison at a false date a year earlier is also non-zero (+0.23, s.e. 0.11). The evidence on junior hiring is also less robust than the evidence on tasks. The software-developer comparison is the only one of ten pre-specified variants with the expected sign that survives the adjustment for multiple testing, and a test across firms of why the dip appeared where it did fails (Appendix D.4).

## 9 Conclusion

This paper asked how firms decide where each technical task is done, and what generative AI changed in that decision and for the workers who do the tasks. Following the task approach, which describes a job by what the worker does (Autor et al. 2003; Acemoglu and Autor 2011), it reads the task from the duties in the job descriptions that the same firms post in five countries. Take two jobs that a firm posts in the same year, one building a model or a data system and one applying it. Between 2019 and 2025 India gained every kind of technical work and its share of these firms' U.S. and Indian technical jobs rose from about one in five to close to one in two, yet the building job's odds of being posted in India rather than the United States stayed about a fifth lower than the applying job's (18 percent). Firms new to

India show the same gap in their first year as several years later. The gap is larger in the Philippines (37 percent) and Mexico (26 percent), and close to zero in the United Kingdom on the observed labels.

The division is not a single ladder. In Grossman and Rossi-Hansberg (2008) one index of the cost of distance ranks all tasks, so every destination takes them in the same order. Here each country takes its own mix of tasks, India technical application, the Philippines functional application and the United Kingdom product and testing, and a formal test rejects a single ranking. The scarcity of doctoral skill accounts for part of the pattern, as the theory that assigns skill-intensive tasks to skill-abundant countries predicts (Costinot 2009; Costinot and Vogel 2010): the work that asks most often for a doctorate is placed least readily where doctorates are scarce, most clearly in the Philippines and Mexico but not in India. Much as German multinationals that place more work in low-income countries keep non-routine and interactive tasks in Germany (Becker et al. 2013), these firms place the work that needs the scarcest skills where those skills are. For U.S. workers in the occupations of these tasks, pay kept pace with other white-collar pay, as the model predicts when the productivity effect of placing tasks in other countries is at least as large as the labor-supply effect (Grossman and Rossi-Hansberg 2008).

Generative AI changed which tasks firms hire for more than where they place them. After ChatGPT's release, U.S. postings for software developers who code applications fell relative to general tech, while postings for those who build models and data systems show no lasting decline. The studies that find weaker demand for early-career workers in exposed U.S. occupations measure exposure by occupation (Westby et al. 2026; Brynjolfsson et al. 2025a; Hosseini and Lichtinger 2025). But one occupation code, Software Developers, holds both the work that declined and the work that did not, and the same application work fell in other occupations too: the decline follows the task, as the variation in tasks within occupations that Autor and Handel (2013) document would lead one to expect. The fall in junior software postings that Westby et al. (2026) document appears in the firms' U.S. offices as a separate and briefer adjustment, a dip in 2023 that had recovered by 2024, which is the temporary freeze that Friebel et al. (2026) predict when AI raises productivity. Where firms place the tasks shows no detectable shift toward India, although the intervals allow moderate shifts, while the same test picks up the large shifts of the move to remote work. That is the pattern the model gives a technology that displaces labor from the tasks it performs wherever they are done (Acemoglu and Restrepo 2019), rather than one that makes distance cheaper.

For technical workers the lesson is that where work is done and how much of it is done are separate margins, and both move at the level of the task. The occupation, the unit at which

labor statistics and most measures of exposure to AI are built, hides both. Some questions remain open. Linked worker records would show whether falls in postings were falls in hires, dated openings of firms' Indian centres would give the division of work a within-firm shock, and Indian pay for the same jobs would show how the gains are shared outside the United States.

# References


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

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.

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

Amiti, M. and Wei, S.-J. (2009). Service offshoring and productivity: Evidence from the US. *The World Economy*, 32(2):203–220. Earlier version: NBER Working Paper 11926, 2006.

Antràs, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–2564.

Antràs, P., Garicano, L., and Rossi-Hansberg, E. (2006). Offshoring in a knowledge economy. *Quarterly Journal of Economics*, 121(1):31–77.

Antràs, P. and Rossi-Hansberg, E. (2009). Organizations and trade. *Annual Review of Economics*, 1:43–64.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Arora, A., Arunachalam, V. S., Asundi, J., and Fernandes, R. (2001). The Indian software services industry. *Research Policy*, 30(8):1267–1287.

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

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

Autor, D., Dube, A., and McGrew, A. (2023). The unexpected compression: Competition at work in the low wage labor market. NBER Working Paper 31010, National Bureau of Economic Research.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, 29(3):3–30.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–1597.

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

Autor, D. H. and Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31(S1):S59–S96.

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

Baldwin, R. (2016). *The Great Convergence: Information Technology and the New Globalization*. Harvard University Press, Cambridge, MA.

Becker, S. O., Ekholm, K., and Muendler, M.-A. (2013). Offshoring and the onshore composition of tasks and skills. *Journal of International Economics*, 90(1):91–106.

Blinder, A. S. (2006). Offshoring: The next industrial revolution? *Foreign Affairs*, 85(2):113–128.

Blinder, A. S. (2009). How many US jobs might be offshorable? *World Economics*, 10(2):41–78.

Blinder, A. S. and Krueger, A. B. (2013). Alternative measures of offshorability: A survey approach. *Journal of Labor Economics*, 31(2, Part 2):S97–S128.

Bornstein, M., Appenzeller, G., and Casado, M. (2023). Who owns the generative AI platform? Andreessen Horowitz. January 19, 2023.

Brynjolfsson, E., Chandar, B., and Chen, R. (2025a). Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence. Working paper, Stanford Digital Economy Lab, November 2025.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2025b). Generative AI at work. *Quarterly Journal of Economics*, 140(2):889–942.

Chandar, B. and Klein Teeselink, B. (2026). How does AI change labor demand? Evidence from 41 countries. Stanford University and King's College London, September 2026.

Costinot, A. (2009). An elementary theory of comparative advantage. *Econometrica*, 77(4):1165–1192.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What goods do countries trade? A quantitative exploration of Ricardo's ideas. *Review of Economic Studies*, 79(2):581–608.

Costinot, A. and Vogel, J. (2010). Matching and inequality in the world economy. *Journal of Political Economy*, 118(4):747–786.

Crinò, R. (2010). Service offshoring and white-collar employment. *Review of Economic Studies*, 77(2):595–632.

Dekle, R., Eaton, J., and Kortum, S. (2007). Unbalanced trade. *American Economic Review*, 97(2):351–355.

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

Dingel, J. I. and Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, 189:104235.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Ebenstein, A., Harrison, A., McMillan, M., and Phillips, S. (2014). Estimating the impact of trade and offshoring on American workers using the Current Population Surveys. *Review of Economics and Statistics*, 96(4):581–595.

Eckert, F., Ganapati, S., and Walsh, C. (2022). Urban-biased growth: A macroeconomic analysis. Opportunity and Inclusive Growth Institute Working Paper 25, Federal Reserve Bank of Minneapolis.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308.

Feenstra, R. C. and Hanson, G. H. (1996). Globalization, outsourcing, and wage inequality. *American Economic Review Papers and Proceedings*, 86(2):240–245.

Felten, E., Raj, M., and Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217.

Friebel, G., Huang, Y., Li, J., Shukla, S., and Zhang, A. (2026). Pyramids, diamonds, and oscillations: AI and the structure of internal labor markets. Working paper, April 10, 2026.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904.

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *Quarterly Journal of Economics*, 121(4):1383–1435.

Glennon, B. (2024). How do restrictions on high-skilled immigration affect offshoring? Evidence from the H-1B program. *Management Science*, 70(2):907–930.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–2526.

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

Grossman, G. M. and Rossi-Hansberg, E. (2012). Task trade between similar countries. *Econometrica*, 80(2):593–629.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4, pages 131–195. Elsevier.

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

Hosseini, S. M. and Lichtinger, G. (2025). Generative AI as seniority-biased technological change: Evidence from U.S. résumé and job posting data. Working paper, SSRN 5425555.

Humlum, A. and Vestergaard, E. (2025). Large language models, small labor market effects. NBER Working Paper 33777.

Hummels, D., Jørgensen, R., Munch, J., and Xiang, C. (2014). The wage effects of offshoring: Evidence from Danish matched worker-firm data. *American Economic Review*, 104(6):1597–1629.

Hummels, D., Munch, J. R., and Xiang, C. (2018). Offshoring and labor markets. *Journal of Economic Literature*, 56(3):981–1028.

Keller, W. and Yeaple, S. R. (2013). The gravity of knowledge. *American Economic Review*, 103(4):1414–1444.

Lewin, A. Y., Massini, S., and Peeters, C. (2009). Why are companies offshoring innovation? The emerging global race for talent. *Journal of International Business Studies*, 40(6):901–925.

Liu, R. and Trefler, D. (2019). A sorted tale of globalization: White collar jobs and the rise of service offshoring. *Journal of International Economics*, 118:105–122.

Nasscom and Zinnov (2024). India GCC landscape report: The 5 year journey. Technical report, Nasscom.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Oldenski, L. (2012). Export versus FDI and the communication of complex information. *Journal of International Economics*, 87(2):312–322.

Ottaviano, G. I. P., Peri, G., and Wright, G. C. (2013). Immigration, offshoring, and American jobs. *American Economic Review*, 103(5):1925–1959.

Pandit, H. and Nambiar, M. (2026). Who builds AI? The global supply of AI workers, migration, gender, and the firms that employ them. Working paper, Northeastern University and Boston College.

Ravanilla, N. (2026). Generative AI and labor displacement in the global south: Evidence from Philippine business process outsourcing. SSRN Working Paper 7313383, August 5, 2026.

Rodríguez-Clare, A. (2010). Offshoring in a Ricardian world. *American Economic Journal: Macroeconomics*, 2(2):227–258.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4):641–658.

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., and Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, pages 2503–2511.

Shukla, S. (2026). Technology bottlenecks: Measuring where AI could raise the value of work it cannot yet do. SSRN Working Paper 7605598, October 21, 2026.

Weidner, M. and Zylkin, T. (2021). Bias and consistency in three-way gravity models. *Journal of International Economics*, 132:103513.

Westby, S., Modestino, A. S., and Cheng, P. (2026). The impact of generative AI on job opportunities for junior software developers. Submitted, Information Systems Research, April 2026.

Xiao, S., Liu, Z., Zhang, P., Muennighoff, N., Lian, D., and Nie, J.-Y. (2023). C-Pack: Packaged resources to advance general Chinese embedding. arXiv:2309.07597.

# A The model in full

This appendix sets out the model that Section 5 describes in words: the assumptions one at a time, the share of each task's jobs placed in each country, the gravity equation and the five propositions tested in Sections 6–8. The mapping from the model to each estimate is in Appendix B and the proofs in Appendix C.

## A.1 The model

A firm $i$ produces output by combining tasks $z \in \mathcal{Z}$—for the bank, the seven jobs of Section 5. The United States is the reference location against which every destination is compared; nothing in the model depends on where the firm is headquartered. In year $t$ the firm has $N_{izt}$ jobs of task $z$ to fill, and it places each job $j$ in one country $c \in \mathcal{C} = \{\text{US}\} \cup \mathcal{D}$, with destinations $\mathcal{D} = \{\text{IN}, \text{PH}, \text{MX}, \text{UK}\}$. Each country has graduate ($H$) and non-graduate ($L$) technical labor with wages $w_c^H$ and $w_c^L$; $\rho_c \equiv \log(w_c^H / w_c^L)$ is the log graduate premium. Task $z$ requires $a(z)$ units of labor per job, a share $s_z \in [0, 1]$ of it graduate labor.

Outside the United States wages are lower, but doing a task far from the firm's U.S. teams costs extra: instructions must be written down, work must be checked from a distance, and some knowledge does not travel. The cost of job $j$ of task $z$ in country $c$ puts these together as in Equation (1) of Section 5, where $\tau_c(z) \geq 1$ is the extra cost of performing $z$ in $c$ relative to the United States ($\tau_{US}(z) = 1$), $\phi_{ic}$ collects everything specific to the firm in the country (its pay premium, the cost of running an office there), and $\varepsilon_{ijc}$ is a job-specific efficiency draw. Each job is placed where its cost is lowest. Put numbers on the bank. Suppose each job costs \$100 in the United States and \$30 in wages in India. The phone app is easy to specify and check, so for it $\tau_{IN} = 1.5$: done from Bengaluru it costs \$45, and it goes to India. Training the model needs the bank's customer data, compute and daily contact with its research team, so $\tau_{IN} = 4$: it would cost \$120 from Bengaluru, and it stays in New York. Reviewing the model for risk sits in between; it may stay in New York in one bank and move to Bengaluru in another.

GRH's original model is the special case with one destination, one kind of labor, a single index of how hard each task is to do at a distance, and no job-specific draws. Line the tasks up from cheapest to send to most expensive, and the firm places every task below a cutoff outside the United States; when distance gets cheaper—better video calls, cloud tools, the shift to remote work in 2020—the cutoff moves up (Appendix C). Today's technical work departs from this case in several directions, and the assumptions below build them in one at a time.

To begin with, a firm places jobs, not whole tasks. Two of the bank's testing jobs can

differ: one team needs someone in its time zone, another has spare desks in Bengaluru. These reasons are invisible to the researcher, and $\varepsilon_{ijc}$ stands for them.

> *Assumption 1 (efficiency draws).* $\varepsilon_{ijc}$ is independent across jobs and countries and distributed Fréchet, $\Pr(\varepsilon \le e) = \exp(-e^{-\theta})$, $\theta > 0$.

The Fréchet distribution is the one Eaton and Kortum (2002) use for trade, and it turns "each job goes where it is cheapest" into a simple formula for the share of a task's jobs in each country. When $\theta$ is large, small cost differences move many jobs, and in the limit every job of a task goes to the cheapest country, as in GRH's cutoff. When $\theta$ is small, jobs of the same task are spread across countries even when one is cheaper. That spread is what the data show: the same title is advertised in several countries in the same year.

Firms also use several destinations, and each has its own cost of distance. A business-analysis job that needs fluent English and contact with British clients is cheap to place in London and costly in Manila; an engineering job may be the reverse.

> *Assumption 2 (location costs).* $\log \tau_c(z) = \kappa_c + x_z' \delta_c$, where $x_z$ is a vector of task characteristics and $\delta_{US} = 0$, $\kappa_{US} = 0$. In the main estimates $x_z$ is the task's category; Appendix M uses measured requirements.

The extra cost has a part common to all of a destination's work ($\kappa_c$, the cost of having an office there at all) and a part that depends on the task's characteristics, with weights $\delta_c$ that differ by destination. That last point is the extension. In GRH one index $t(z)$ ranks tasks and only the scale $b$ differs across destinations, so every destination ranks the tasks in the same order; here each destination can have its own order.

Then there is skill. The bank's model training needs PhDs; its dashboards do not. Countries differ in how many graduates, and especially doctoral graduates, they have, so graduate-heavy work is expensive to place where graduates are scarce. The next pair of assumptions makes this precise.

> *Assumption 3 (task skill intensity).* $s_z$ does not vary across countries.
>
> *Assumption 4 (graduate premia).* In each country the relative demand for graduate technical labor has constant elasticity $\sigma > 0$: $H_c/L_c = B_c\, e^{-\sigma \rho_c}$, so $\rho_c = \sigma^{-1}\big(\log B_c - \log(H_c/L_c)\big)$. The firms take $\rho_c$ as given.

Assumption 3 says that model training uses the same share of graduates in Bengaluru as in New York. It rules out a firm redesigning a task to use fewer graduates where they are scarce; if firms do that, part of the skill cost is avoided, and the skill coefficients estimated

in Section 6.3 understate how much graduate scarcity matters. Assumption 4 is the simplest link between supply and pay: where graduates are scarcer relative to demand, the premium $\rho_c$ is higher, and any one firm is too small to move it. The assumption is used only to sign how the skill coefficients differ across destinations (Proposition 2); the gravity equation does not need it.

## A.2 Where each job goes

With these pieces, the bank's choice for each job becomes a share: a country receives more of a task's jobs the cheaper it is for that task, and $\theta$ sets how sharply the jobs respond.

*Lemma 1 (job shares).* Under Assumption 1 the share of firm $i$'s jobs of task $z$ placed in country $c$ is

$$\pi_{izc} \;=\; \frac{\left(\tau_c(z)\,(w_c^H)^{s_z}(w_c^L)^{1-s_z}\,\phi_{ic}\right)^{-\theta}}{\sum_{c'}\left(\tau_{c'}(z)\,(w_{c'}^H)^{s_z}(w_{c'}^L)^{1-s_z}\,\phi_{ic'}\right)^{-\theta}}, \tag{7}$$

and under Assumptions 1–3 the log odds of destination $d$ against the United States are

$$\frac{1}{\theta}\log\frac{\pi_{izd}}{\pi_{iz,US}} \;=\; \underbrace{\omega_{id}-\kappa_d}_{\text{common to the firm's tasks}} \;-\; \underbrace{x_z'\delta_d}_{\text{location cost}} \;-\; \underbrace{s_z\,(\rho_d-\rho_{US})}_{\text{skill cost}}, \tag{8}$$

where $\omega_{id} \equiv \log\left(w_{US}^L\phi_{i,US}/(w_d^L\phi_{id})\right)$ is the wage wedge.

Read Equation (8) through the bank. India's lower wages ($\omega_{id}$) and the cost of the bank's Bengaluru office ($\kappa_d$) are the same for the app and for model training. They decide how much of its work the bank places in India, but not which jobs move, because they shift every task's odds by the same amount. Which jobs go is decided by the two terms that differ across tasks: model training's location cost $x_z'\delta_d$, the \$120 against \$45 of the example, and its skill cost $s_z(\rho_d-\rho_{US})$, high because it needs PhDs where PhDs are scarce. The size of a task, $a(z)$, is the same in every country and cancels: model training stays in New York not because it is large or important but because it costs relatively more to do elsewhere.

The data see only a fraction $\eta_{ict}$ of the firm's postings in a country and year (LinkUp's coverage), so Lemma 1 is taken to the data as a count model. The model that results is a gravity equation, the workhorse of international trade, and the analogy is close. Think of the bank as a buyer of tasks and of each country as a seller. How much cloth the United States imports from Vietnam depends on how much cloth it buys in all, on how competitive Vietnam is across everything it sells, and on how good Vietnam is at making cloth in particular. The

bank's hiring is governed by the same forces. The number of model-training jobs it places in India depends on how many such jobs it needs, on how attractive India is for all of its work—because of wages there, the size of its Bengaluru office or how fully LinkUp sees its Indian postings—and on how costly model training in particular is to do from India. Fixed effects absorb the bank's overall need and India's general attractiveness, so the estimates isolate the cost of the particular task. It is the task-level counterpart of the country × industry term that Costinot et al. (2012) estimate for trade in goods and, as in their work, it is identified from a double comparison: two kinds of work in two countries, within one firm in one year (Table 4).

Proposition 1 (Gravity equation for jobs). *Under Assumptions 1–3, and if coverage $\eta_{ict}$ does not vary across the firm's tasks within a country and year, the expected number of observed jobs of task $z$ that firm $i$ places in country $c$ in year $t$ is*

$$\mathbb{E}[V_{iczt}] = \exp\big(\alpha_{izt} + \gamma_{ict} + x_z'\beta_c^{\tau} + s_z\,\beta_{ct}^{s}\big), \qquad \beta_c^{\tau} = -\theta\,\delta_c, \quad \beta_{ct}^{s} = -\theta\,(\rho_{ct} - \rho_{US,t}), \tag{9}$$

*where $\alpha_{izt}$ varies only by firm, task and year and $\gamma_{ict}$ only by firm, country and year.*

Neither fixed effect has to be measured. What is estimated, and what the paper is about, is what remains: the destination-specific prices of the task characteristics (in the main estimates, the categories), $\beta_d^{\tau}$, and of graduate intensity, $\beta_d^{s}$, both scaled by $\theta$. Section 6 explains each term; Appendix C shows why Poisson pseudo-maximum likelihood with these fixed effects estimates them consistently.

## A.3 What the model predicts

The predictions follow from Lemma 1 and Proposition 1. Every destination saves the bank money on the non-graduate part of a job, but only a destination with graduates at least as plentiful as in the United States also saves on the graduate part. The more of a job is graduate work, the less a graduate-scarce destination has to offer it. Across destinations, this becomes a ranking.

Proposition 2 (Skill sorting). *Hold graduate demand fixed ($B_d = B_{US}$). (i) The odds $\pi_{izd}/\pi_{iz,US}$ fall with skill intensity $s_z$ if and only if graduates are relatively scarcer in $d$ than in the United States, $H_d/L_d < H_{US}/L_{US}$. (ii) The skill coefficients are ordered by relative graduate supply: $\beta_d^s > \beta_{d'}^s$ if and only if $H_d/L_d > H_{d'}/L_{d'}$. Equivalently, the odds of placing a task in a destination are log-supermodular in the task's graduate intensity and the destination's relative graduate supply.*

Section 6.3 sets out the ordering this implies across the four destinations.

The bank keeps model training in New York because it needs PhDs and because it needs the bank's data, compute and research team close by. The need for PhDs fades where PhDs are plentiful; the need for proximity need not, and it need not be the same in London as in Bengaluru. Differencing Equation (8) between a Building task and an Application task in the same destination removes $\omega_{id}$ and $\kappa_d$ and separates the two needs: a difference in location attributes plus a Building-specific location cost $\delta_{d,b,t}$ that the measured attributes miss, and a difference in skill intensity times the destination's premium gap (Equation (23) in the appendix).

PROPOSITION 3 (What stays in the United States, and when it moves). *Write the Building–Application gap in placement odds as $\Delta_d = \mathcal{L}_d + \mathcal{S}_d$, the sum of a location component $\mathcal{L}_d = -\theta[(x_{z_B} - x_{z_A})'\delta_d + \delta_{d,b,t}]$ and a skill component $\mathcal{S}_d = -\theta(s_{z_B} - s_{z_A})(\rho_{dt} - \rho_{US,t})$. Hold graduate demand fixed as in Proposition 2, and let Building be more graduate-intensive than Application. (i) The skill component is negative where graduates are relatively scarcer than in the United States, more negative the scarcer they are, and zero where relative supply equals that of the United States. The total gap $\Delta_d$ has the same ordering across destinations if the location component is the same in every destination; otherwise destinations are ordered by $\mathcal{L}_d + \mathcal{S}_d$, and where graduate supply equals that of the United States the gap equals $\mathcal{L}_d$. (ii) Once destination × skill intensity is controlled, the skill component disappears and the remaining destination × Building coefficient measures $\mathcal{L}_d$. (iii) The gap changes over time only through the Building-specific location cost or the destination's relative graduate supply. In particular, a fall in location costs common to every task above routine support, $\Delta \log \tau_d(z) = -\xi_{dt}$ for all such $z$, raises the destination's odds for every such task relative to support by $\theta\xi_{dt}$ and leaves the gap unchanged.*

So the gap need not vanish where graduates are plentiful; only its skill part does, and part (ii) says how to separate the two parts in the data. Part (iii) gives a precise meaning to a destination *learning its way up the stack*: if Bengaluru gets cheaper for everything the bank does above keeping the servers running, it gains the app, the dashboards and model training at once, and the gap at the top closes only if India's graduate supply rises or the specific cost of model training falls (P3, P5).

## A.4 Technical change: relocation versus location-neutral labor saving

Generative AI can enter the bank's decision through either part of the cost. In one story AI makes distance cheaper: it writes the documentation, reviews the code, summarizes the

meeting. That lowers the extra cost of every job done outside the United States, as remote work did. In the other, AI writes code and saves the same junior hours wherever the code is written. If a coding assistant cuts the phone app's labor by 30 percent in both countries, the app costs $70 in the United States and $31.50 in India. India is still cheaper by exactly the same ratio, so the job does not move; the bank simply hires fewer juniors in both places. Formally, let each job of task $z$ use junior labor in proportion $j_z$ of $a(z)$, and write $\hat{x} \equiv x'/x$ for the proportional change in any variable, as in exact hat algebra (Dekle et al. 2007). The relocation story acts on Equation (1) through location costs $\hat{\tau}_c(z)$, the labor-saving story through the labor requirement $\hat{a}(z)$.

Proposition 4 (Two channels for generative AI: relocation and location-neutral labor saving). *(i) If $\hat{\tau}_d(z) < 1$ for all $z$ in a destination $d$, that destination's share of every task rises and the U.S. share falls, $\hat{\pi}_{izd} > 1 > \hat{\pi}_{iz,US}$, and junior hiring in $d$ rises relative to the United States. (ii) If AI lowers the junior labor requirement of task $z$ by a factor $1 - \psi_z$ in every country, then $\hat{\pi}_{izc} = 1$ for all $c$: no job moves, and junior hiring in country $c$ falls by*

$$\Delta \log J_c \;\approx\; -\sum_z \nu_{zc}\, \psi_z, \tag{10}$$

*where $\nu_{zc}$ is task $z$'s share of junior hiring in country $c$.*

Part (ii) is the $70 and $31.50 of the app: a saving common to all locations is, like $a(z)$ itself, absent from Equation (7). The two parts therefore make opposite predictions about the same data: under (i) a destination's share of every task jumps when the technology arrives, under (ii) no share moves and only junior hiring falls. Inside the firm, only the part of a jump that differs across tasks can be seen. A fall in location costs that is the same for every task raises the odds of general tech as much as those of any other category, and the firm × country × quarter effect absorbs it together with changes in coverage. The test of Section 8.1 therefore asks whether the technology lowered the cost of distance more for some tasks than for others, as tools that make codified work easier to direct and check would. Under (ii) the incidence across countries is set by the allocation. The codified, low-$s_z$ slice of each occupation is already placed in India, so the junior decline is broad there and concentrated in the most exposed occupation in the United States; an occupation-level contrast therefore detects the effect in the United States more easily than in India (P6).

## A.5 U.S. wages

What happens to the bank's U.S. workers when the app goes to Bengaluru? GRH show that forces pull both ways. Moving the app to Bengaluru saves the bank money, which works

like a rise in productivity: U.S. workers in the remaining tasks become more valuable (the *productivity effect*). The workers who used to build the app now compete for other jobs, which pushes wages down (the *labor-supply effect*). And cheaper production can lower the price of what the bank sells (the *relative-price effect*). To isolate the productivity effect, group tasks into Building and Application, let $w_g$ be the wage of the U.S. workers who perform group $g$, and shut the relative-price effect off.

> *Assumption 5 (small open economy).* The price $p_g$ of the output of group $g$ is set in world markets and firms earn zero profits. Within a group, tasks enter with Cobb–Douglas cost shares $\lambda_z$, $\sum_{z \in g} \lambda_z = 1$.

The bank cannot raise its prices when its costs fall, so any saving from placing jobs outside the United States shows up as higher pay for the U.S. workers who remain, as a better machine would; within a group each task takes a fixed share of costs.

Proposition 5 (Productivity effect). *Under Assumptions 1–5, with labor requirements and U.S. office costs $\phi_{i,US}$ fixed,*

$$\Delta \log w_g \;=\; -\frac{1}{\theta} \sum_{z \in g} \lambda_z \, \Delta \log \pi_{z,US} \;+\; \Delta \log p_g. \tag{11}$$

The gain depends only on how far the U.S. share of the group's jobs has fallen and on $\theta$. When $\theta$ is large, a job's costs differ little across countries, so a job that moves out of the United States saves little and each point of U.S. share is worth less. The U.S. share is a sufficient statistic, as the domestic expenditure share is for the gains from trade in Arkolakis et al. (2012).

The labor-supply and relative-price effects enter through $\Delta \log p_g$. If moving the tasks lowers the price of the group's output by more than the cost saving, pay in the group falls. The test is therefore one-sided: if the productivity effect dominates, pay in the group whose work moved does not fall behind comparable work (P7, Section 7). Appendix B closes the model with one U.S. labor market for white-collar workers, who move between Building, Application and other professional work (Proposition 7); the gain is then shared by all of them, and pay in Building and Application rises alike (P7′).

# B Model details and the mapping to estimates

This appendix collects the parts of the model that the main text summarizes: how the extension differs from GRH, the role of migration, the common white-collar wage, the experience ladder, and what each estimated coefficient measures.

## B.1 What the extension adds

GRH model the placement of tasks across countries as trade in tasks and show that a fall in the cost of placing tasks in another country can *raise* the wage of the U.S. workers who remain, because it acts like factor-augmenting technical progress. To make that point they assume one foreign country, homogeneous labor, a single index $t(z)$ that ranks tasks by their cost of remote performance, and all-or-nothing allocation of each task. Table 12 sets their environment beside the one used here.

Table 12: The GRH environment and the extension

| | Grossman and Rossi-Hansberg (2008) | This paper |
|---|---|---|
| Unit of allocation | A task, all or nothing | A job; a task's jobs are split across countries by idiosyncratic cost draws |
| Destinations | One | Several, each with its own costs |
| Cost of remote performance | One index $t(z)$: all destinations rank tasks alike | $\log \tau_c(z) = \kappa_c + x_z' \delta_c$: several sources, priced differently by each destination |
| Labor | Homogeneous | Graduate and non-graduate; tasks differ in graduate intensity, destinations in graduate supply; graduates migrate |
| Technical change | A fall in the cost of placing tasks in another country | A fall in location costs $\tau$, or a fall in labor requirements $a$ common to all locations |
| U.S. wages | Productivity, relative-price and labor-supply effects | Productivity effect with the U.S. share as a sufficient statistic; with one white-collar labor market, moving tasks out of the United States raises the common wage |
| Empirical content | Comparative statics | A gravity equation for jobs with firm fixed effects, estimable from postings |

The extension delivers results that the GRH environment cannot.

- *Destination-specific rankings.* The same task can be cheap to perform in the United Kingdom and costly in the Philippines. GRH's single ranking is the restriction $\delta_c = \delta$ for all $c$; destinations in fact order the seven categories very differently (Section 6.2).

- *Skill sorting.* The relative odds of placing a task in a destination are log-supermodular in the task's graduate intensity and the destination's relative graduate supply, so graduate-intensive tasks stay in the United States where graduates are scarce. Graduate emigration to the United States steepens the gradient (Propositions 2 and 6).

- *Climbing the stack without closing the gap at the top.* A fall in location costs common to all tasks above routine support raises the destination's odds for every such task and

leaves the Building–Application gap unchanged (Proposition 3).

- *Neutrality of labor-saving technology.* Labor requirements that are common across locations cancel from the allocation, so a technology that saves the same labor wherever a task is performed moves no tasks; only a fall in location costs does (Proposition 4).

- *A sufficient statistic for the productivity effect.* With world output prices, the gain to U.S. workers is $1/\theta$ times the fall in the log U.S. share of jobs, the task-trade analogue of Arkolakis et al. (2012). With one white-collar labor market it accrues to all white-collar workers (Propositions 5 and 7).

## B.2 Migration and graduate supply

Graduate supply in a destination is itself endogenous to migration. Let $\bar{H}_d$ be the graduates trained in $d$ and $m_d$ the share who work in the United States, so $H_d = (1 - m_d)\bar{H}_d$ and $H_{US} = \bar{H}_{US} + \sum_d m_d \bar{H}_d$. India keeps only 37 percent of its graduate-educated AI workers, and half of them live in the United States (Pandit and Nambiar 2026).

Proposition 6 (Migration and the location of graduate-intensive tasks). *Holding $\bar{H}$, $L$ and $B$ fixed, a rise in $m_d$ lowers $\beta^s_d$, and lowers $\beta^s_{d'}$ for every other destination $d'$ through the rise in $H_{US}$.*

In words: when an Indian graduate moves to the United States, India has one graduate fewer and the United States one more. Both changes widen the gap in graduate supply, so graduate-intensive tasks become less attractive to place in India, and in every other destination too, because each is compared with the United States. Hiring the migrant and placing the task outside the United States are therefore substitutes for graduate-intensive work, so restricting high-skilled immigration moves graduate-intensive tasks out of the United States, as Glennon (2024) finds for the H-1B program. I do not observe migration by firm, and the proposition organizes facts rather than being tested.

## B.3 One white-collar labor market

Engineers, analysts and managers move between Building and Application work and between them and other white-collar work. Closing the model with a common wage makes the productivity effect economy-wide.

> *Assumption 6 (one white-collar labor market).* U.S. white-collar workers are mobile across task groups and earn a common wage $w_W$. The groups combine

into traded white-collar services with Cobb–Douglas cost shares $\mu_g$, $\sum_g \mu_g = 1$, whose price $p$ is set in world markets.

In words: an engineer doing Building work in the United States could instead do Application work or another white-collar job, so pay in the three cannot drift apart for long. The assumption fails if Building skills are specific enough that its workers cannot move; Proposition 5 then applies group by group.

PROPOSITION 7 (Task placement and the U.S. white-collar wage). *Under Assumptions 1–3, 5 and 6,*

$$\Delta \log w_W \;=\; \frac{1}{\theta} \sum_g \mu_g \Big( - \sum_{z \in g} \lambda_z \, \Delta \log \pi_{z,US} \Big) \;+\; \Delta \log p. \tag{12}$$

*(i) With $p$ fixed, moving any group's tasks out of the United States raises the wage of every U.S. white-collar worker, including those in groups whose U.S. share did not fall. (ii) Pay in different task groups moves together. (iii) The wage falls only if moving them lowers the world price of white-collar services by more than $1/\theta$ times the cost-weighted fall in U.S. shares.*

In words: when any part of white-collar work becomes cheaper by moving out of the United States, white-collar services as a whole become cheaper to produce, and with their price fixed the saving goes to the one wage that all white-collar workers earn. This is GRH's small-open-economy logic, in which goods prices pin down factor prices, applied to one labor market. A Building engineer in New Jersey gains when testing work moves to India for the same reason a worker gains from technical progress in a complementary task. The proposition also disciplines the evidence. Part (ii) predicts equal pay growth in Building and Application, and it implies that comparisons with other mobile white-collar work difference out the gain: such comparisons test the labor-supply effect, not the size of the productivity effect (P7′).

## B.4 The experience ladder within each country

Equation (10) holds senior hiring fixed. To describe the whole ladder, give each country's operation a three-layer knowledge hierarchy (Garicano and Rossi-Hansberg 2006). Juniors execute codified tasks; mid-level engineers execute and solve the problems within their knowledge $k_c$, a share $F(k_c)$; the rest are escalated to seniors, each taking $h$ units of senior time. With $\ell_c$ juniors per mid-level engineer,

$$n_c^J : n_c^M : n_c^S \;=\; \ell_c \;:\; 1 \;:\; q_c, \qquad q_c \equiv h\,[1 - F(k_c)]. \tag{13}$$

In a team that spans countries (Antràs et al. 2006), problems that need tied or doctoral knowledge are escalated to the United States, so seniors outside the United States mostly resolve local exceptions. Generative AI lowers $\ell_c$ by performing codified tasks and lowers $q_c$ by putting knowledge within reach.

PROPOSITION 8 (The shape of the ladder). *Let AI lower $\ell_c$ by the proportion $\hat{\ell}_c$ and $q_c$ by the proportion $\hat{q}_c$. (i) The mid-level share rises. (ii) If $\hat{q}_c = 0$, the junior share falls and the senior share rises. (iii) Junior and senior shares both fall—the ladder becomes a diamond—if and only if*

$$\frac{\ell_c}{1+\ell_c} \;<\; \frac{\hat{q}_c}{\hat{\ell}_c} \;<\; \frac{1+q_c}{q_c}.$$

In words: AI does two things to a team. It does some of the codified work juniors did, so fewer juniors are needed per mid-level engineer, and it lets mid-level engineers solve problems they used to pass up, so fewer seniors are needed. Mid-level engineers gain share either way, because AI makes them more capable rather than replacing them. Whether the bottom or the top shrinks as well depends on which effect is larger. If AI only replaces junior work, seniors gain share. If it also takes over enough of what seniors did—more than the first inequality requires, but not so much that juniors gain share, which is the second—both ends shrink and the pyramid becomes a diamond. Table 15 in Appendix D describes the two ladders; it is a description, not a test of the proposition.

The selection of codified work into India makes both margins large there: junior work there is mostly codified, and seniors there resolve problems that AI puts within reach of mid-level engineers. In the United States, seniors resolve frontier and doctoral problems that AI reaches least, so $\hat{q}_{US}$ is small and part (ii) applies. The proposition is static; Friebel et al. (2026) supply the dynamics, in which a productivity shock causes a temporary junior freeze and a learning shock a permanent change of shape (P6′).

## B.5 What each coefficient measures

The outcome is the number of distinct jobs, not postings. A distinct job is a firm × country × title × creation date, labelled by its earliest posting (Section 3.3). For firm $i$, country $c$, category $k$ and year $t$, $V_{ickt}$ counts these jobs, with a zero for every country and category in which a firm that posts in year $t$ places none. Distinct jobs are the model's $j$: LinkUp writes one posting per location, and it does so more often for some categories and countries than for others, so posting counts would weight the same job differently by country (Appendix N).

Let $\mathcal{Z}_k$ be the tasks in category $k$ and suppose skill intensity is common within a category,

$s_z = \bar{s}_k$. Summing Equation (9) over $z \in \mathcal{Z}_k$, the category count has the same form,

$$\mathbb{E}[V_{ickt}] = \exp\big(\alpha_{ikt} + \gamma_{ict} + \beta_{c,k}\big), \qquad \alpha_{ikt} = \log \textstyle\sum_{z \in \mathcal{Z}_k} e^{\alpha_{izt}}, \quad \beta_{c,k} = -\theta\big[\delta_{c,k} + \bar{s}_k(\rho_c - \rho_{US})\big], \tag{14}$$

where $\delta_{c,k}$ is the location cost of category $k$ in $c$ ($\delta_{US,k} = 0$). When $s_z$ varies within a category, $\bar{s}_k$ is replaced by a job-weighted average and the equation holds approximately. The firm × category × year effect absorbs the firm's demand for the category, and the firm × country × year effect absorbs anything common to all of the firm's categories in a country, including $\beta_{c,g}$ for the reference category. Only differences from a reference are therefore identified. Each coefficient $b_{d,k}$ of Equation (2) therefore measures

$$b_{d,k} \equiv \beta_{d,k} - \beta_{d,g} = -\theta\Big[\underbrace{(\delta_{d,k} - \delta_{d,g})}_{\text{location cost}} + \underbrace{(\bar{s}_k - \bar{s}_g)(\rho_d - \rho_{US})}_{\text{skill cost}}\Big]. \tag{15}$$

In the data, $b_{d,k}$ is a log odds ratio. Within firm and year,

$$b_{d,k} = \log \frac{\pi_{dk}/\pi_{US,k}}{\pi_{dg}/\pi_{US,g}}, \tag{16}$$

the odds of placing a category-$k$ job in $d$ rather than the United States, relative to the same odds for a general-tech job. A coefficient of 0.22 means odds $e^{0.22} = 1.25$ times those of general tech. Comparisons between destinations follow from the same regression, since the United States cancels: $b_{d,k} - b_{d',k}$ is the log odds of $d$ rather than $d'$.

The Building–Application difference of Equation (3) maps into the model in the same way,

$$\Delta_d = -\theta\Big[(\delta_{d,B} - \delta_{d,A}) + (\bar{s}_B - \bar{s}_A)(\rho_d - \rho_{US})\Big], \tag{17}$$

the empirical counterpart of Equation (23). Proposition 3 predicts that the skill part, $-\theta(\bar{s}_B - \bar{s}_A)(\rho_d - \rho_{US})$, is negative where graduates are scarce and zero where they are as plentiful as in the United States. It places no restriction on the location part $-\theta(\delta_{d,B} - \delta_{d,A})$, which part (ii) isolates by adding the skill term to the regression. Without the skill term, the single ranking $\delta_{d,k} = \lambda_d\, t_k$ gives Equation (4). With it, Equation (15) is a two-factor model and orderings can differ, because $\rho_d - \rho_{US}$ is positive in some destinations and near zero in others. With six categories the rank correlations are descriptive; Appendix M gives a formal test on measured requirements.

Skill intensity is a property of titles, so the skill price is estimated on firm × country × title × year cells (Equation (5)), using the titles of the fixed lookup (Appendix H) and $s_z$

standardized across titles:

$$\beta_d^s = -\theta\,(\rho_d - \rho_{US}) = \frac{\theta}{\sigma}\Big[\log\frac{H_d}{L_d} - \log\frac{H_{US}}{L_{US}}\Big] - \frac{\theta}{\sigma}\log\frac{B_d}{B_{US}}, \tag{18}$$

the second equality from Assumption 4 (Equation (22)). Proposition 2 predicts $\beta_d^s < 0$ where graduates are scarcer than in the United States, ordered by relative supply. Part (ii) of Proposition 3 is tested by adding $s_z\beta_d^s$ to the Building–Application regression on title cells, which removes the skill term from Equation (3).

The gravity coefficients identify products of $\theta$ with prices, not $\theta$ itself. The productivity effect of Equation (11) needs $\theta$ and the change in the U.S. share,

$$\Delta \log w_g\big|_{p_g \text{ fixed}} = -\frac{1}{\theta}\,\Delta \log \pi_{g,US}, \qquad \pi_{g,US} = \frac{\sum_{i,k\in g} V_{i,US,kt}}{\sum_{i,c,k\in g} V_{ickt}}, \tag{19}$$

with $\theta$ between 2.2 and 4.8, values used for illustration because these data do not identify $\theta$ (Section 7). Equation (19) simplifies Equation (11) in two ways. It treats each group as one task: the log change in the group's aggregate U.S. share equals the cost-weighted average of its tasks' log changes only if every task's U.S. share falls by the same proportion. And it measures the U.S. share against U.S. plus Indian jobs, leaving out the other destinations. Section 7 computes the statistic as Equation (11) defines it, task by task with cost weights and all five destinations: $-0.42$ for Application and $-0.48$ for Building, against $-0.37$ for each on the simplified measure, which gives 9–19 and 10–22 log points. Both illustrate a magnitude; neither is an estimate or a bound, because $\theta$ is not identified.

The gravity estimates identify products of $\theta$ with attribute prices or premium gaps, not either alone; they rank and compare destinations but do not price $\tau$ or $\rho$ in dollars. With four destinations the skill-supply prediction is a test of ordering, not an estimate of $\theta/\sigma$. Proposition 5 isolates the productivity effect at fixed prices; its empirical counterpart compares pay across groups and benchmarks and uses the implied $\theta$ only as a consistency check. The allocation estimates describe how firms placed tasks over 2019–2025; the one shock-based test, of Proposition 4, is a test for a break in that allocation.

# C Proofs for the model

This appendix derives the results of Section 5 in the order in which they appear there.

Table 13: Model objects, their empirical counterparts and the estimates

| Object | Model meaning | Estimated from | India | Philippines | Mexico | U.K. |
|---|---|---|---|---|---|---|
| $b_{d,k}$ | Price of category $k$ in $d$ relative to general tech, $-\theta[\Delta\delta + \Delta\bar{s}\,\Delta\rho]$ | Eq. (2), distinct jobs, full text | | Table 5, Panel A | | |
| $\Delta_d$ | Building − Application, Eq. (3) | Eq. (2), grouped | −0.20 | −0.46 | −0.30 | −0.01 |
| | | | (0.04) | (0.05) | (0.04) | (0.03) |
| $1 - e^{\Delta_d}$ | Shortfall in Building's odds | from $\Delta_d$ | 18% | 37% | 26% | 1% |
| $r_{IN,d}$ | Rank correlation with India's ordering; 1 under Eq. (4) | $b_{d,k}$, six categories | – | 0.43 | 0.83 | −0.43 |
| $\beta^s_d$ | Skill price, $-\theta(\rho_d - \rho_{US})$, per s.d. of PhD intensity | title × category cells | −0.10 | −0.35 | −0.23 | −0.10 |
| | | | (0.02) | (0.06) | (0.04) | (0.03) |
| $\Delta_d \mid s_z$ | Gap net of skill: $-\theta(\delta_{d,B} - \delta_{d,A})$ | Grouped, title × category cells, with $s_z\beta^s_d$ | −0.11 | −0.34 | −0.02 | −0.05 |
| | | | (0.05) | (0.08) | (0.07) | (0.05) |
| $\theta$ | Placement elasticity | Not identified; 2.2–4.8 used for illustration | | Section 7 | | |
| $\Delta \log w_g$ | Productivity effect at fixed prices, 2019–2025, Eq. (11) | U.S. share of each title's jobs, five countries, cost weights | 9–19 (10–22) log points; illustrative | | | |

*Notes:* Standard errors clustered by firm in parentheses. $\Delta\delta = \delta_{d,k} - \delta_{d,g}$, $\Delta\bar{s} = \bar{s}_k - \bar{s}_g$, $\Delta\rho = \rho_d - \rho_{US}$. Rows $\Delta_d$ and $r_{IN,d}$: Table 5 (all technical jobs, English-language postings, 2019–2025). Rows $\beta^s_d$ and $\Delta_d \mid s_z$: Table 7, Panels A(1) and C (full-text categories on lookup titles, the titles for which skill intensity is measured; the gap before the skill control is −0.15, −0.42, −0.11 and −0.09 on these titles). The productivity row uses the U.S. share of U.S.-plus-India jobs of Section 7, treats each group as one task, and is an illustration, not a bound; part of the fall in that share reflects LinkUp's expanding coverage of India.

## Shares from job-level draws

By Equation (1), job $j$ of task $z$ at firm $i$ costs $C_{ijzc} = k_{izc}/\varepsilon_{ijc}$ in country $c$, with deterministic part

$$k_{izc} \equiv a(z)\,\tau_c(z)\,(w_c^H)^{s_z}(w_c^L)^{1-s_z}\,\phi_{ic}.$$

For Lemma 1, read $\tau_c(z)(w_c^H)^{s_z}(w_c^L)^{1-s_z}$ as a single unit cost $w_c\,t_c(z)$; Assumptions 2 and 3 give it structure, and the derivation of the shares does not depend on that structure.

Under Assumption 1, for any $x > 0$,

$$\Pr(C_{ijzc} > x) = \Pr\big(\varepsilon_{ijc} < k_{izc}/x\big) = \exp\big(-(x/k_{izc})^{\theta}\big),$$

so the density of $C_{ijzc}$ is $g_c(x) = \theta\,x^{\theta-1}k_{izc}^{-\theta}\exp\big(-(x/k_{izc})^{\theta}\big)$.

The job goes to $c$ if $C_{ijzc}$ is below every other country's cost. With draws independent across countries,

$$\pi_{izc} = \int_0^\infty g_c(x)\prod_{c'\neq c}\Pr(C_{ijzc'} > x)\,dx = \int_0^\infty \theta\,x^{\theta-1}k_{izc}^{-\theta}\exp\Big(-x^{\theta}\textstyle\sum_{c'}k_{izc'}^{-\theta}\Big)\,dx = \dfrac{k_{izc}^{-\theta}}{\sum_{c'}k_{izc'}^{-\theta}},$$

using $\int_0^\infty \theta x^{\theta-1}e^{-Kx^{\theta}}dx = 1/K$. The factor $a(z)^{-\theta}$ appears in the numerator and in every term of the denominator and cancels. With the structure of Assumptions 2 and 3, this is Equation (7).

As $\theta \to \infty$, $\pi_{izc} \to 1$ for the country with the lowest $k_{izc}$ and 0 for the others. GRH's original model is the special case with one destination $F$, one kind of labor ($s_z = 0$), a single difficulty index $\tau_F(z) = b\,t(z)$ and no job-specific draws. With $k_{iz,F}/k_{iz,US} = w_F\,b\,t(z)/w_{US}$, every job of task $z$ is placed in $F$ if and only if

$$w_F\,b\,t(z) < w_{US}, \tag{20}$$

so the firm places every task below a cutoff in $t(z)$ in $F$, and a fall in $b$ moves the cutoff up.

One strong implication of the logit form is that destinations are substitutes in fixed proportions, so a cheaper India draws jobs from London and New York in proportion to their shares. This matters for counterfactuals that move jobs among destinations, which the paper does not compute, and not for the estimates, each of which compares one destination with the United States.

## Location costs, skill costs and the comparison with the United States

Taking logs of $k_{izc}$ and using Assumption 2 and $\rho_c = \log(w_c^H/w_c^L)$,

$$\log k_{izc} = \log a(z) + \kappa_c + x_z'\delta_c + \log w_c^L + s_z\,\rho_c + \log \phi_{ic},$$

since $(w_c^H)^{s_z}(w_c^L)^{1-s_z} = w_c^L\, e^{s_z\rho_c}$. Assumption 3 is what allows $s_z$ to carry no country subscript. Dividing $\pi_{izd}$ by $\pi_{iz,US}$ removes the denominator of the share. With $\kappa_{US} = 0$ and $\delta_{US} = 0$,

$$\log \frac{\pi_{izd}}{\pi_{iz,US}} = -\theta\big(\log k_{izd} - \log k_{iz,US}\big) = \theta\Big[\underbrace{\log \tfrac{w_{US}^L \phi_{i,US}}{w_d^L \phi_{id}}}_{\omega_{id}} - \kappa_d - x_z'\delta_d - s_z(\rho_d - \rho_{US})\Big],$$

which is Equation (8). Assumption 4 states the premium as

$$\rho_c = \tfrac{1}{\sigma}\big(\log B_c - \log(H_c/L_c)\big), \tag{21}$$

and substituting it into $\beta_d^s = -\theta(\rho_d - \rho_{US})$ gives

$$\beta_d^s \;=\; \frac{\theta}{\sigma}\Big[\log \frac{H_d}{L_d} - \log \frac{H_{US}}{L_{US}}\Big] \;-\; \frac{\theta}{\sigma}\big[\log B_d - \log B_{US}\big]. \tag{22}$$

*Proof of Proposition 2.* From Equation (8), $\partial \log(\pi_{izd}/\pi_{iz,US})/\partial s_z = -\theta(\rho_d - \rho_{US})$, which is negative if and only if $\rho_d > \rho_{US}$. By Equation (21) with $B_d = B_{US}$, $\rho_d - \rho_{US} = \sigma^{-1}[\log(H_{US}/L_{US}) - \log(H_d/L_d)]$, positive if and only if $H_d/L_d < H_{US}/L_{US}$. This proves (i). For (ii), Equation (22) with $B_d = B_{d'}$ gives $\beta_d^s - \beta_{d'}^s = (\theta/\sigma)[\log(H_d/L_d) - \log(H_{d'}/L_{d'})]$, which has the sign of $H_d/L_d - H_{d'}/L_{d'}$. □

Because $\theta/\sigma > 0$ multiplies every destination's coefficient, the ordering in (ii) does not depend on the values of $\theta$ and $\sigma$. It would survive measurement error in $s_z$ that attenuated every destination's coefficient by the same factor, but that is not guaranteed (see the proof of Proposition 1). It does depend on the demand shifters $B_d$; a lower $B_d$ raises $\beta_d^s$.

## The gravity equation and its estimation

*Proof of Proposition 1.* With $N_{izt}$ jobs of task $z$ placed independently, the number placed in $c$ is multinomial with mean $N_{izt}\pi_{izct}$. LinkUp observes each posting in $c$ with probability $\eta_{ict}$, independently of the task, so the observed count $V_{iczt}$ has mean $\eta_{ict}N_{izt}\pi_{izct}$. Write

$\Phi_{izt} \equiv \sum_{c'} k_{izc't}^{-\theta}$. Taking logs and substituting $\log k_{izc}$,

$$\log \mathbb{E}[V_{iczt}] = \underbrace{\log N_{izt} - \log \Phi_{izt} - \theta \log a(z)}_{\alpha_{izt}} \underbrace{- \theta \left(\log w_{ct}^{L} + \kappa_{ct} + \log \phi_{ict}\right) + \log \eta_{ict}}_{\gamma_{ict}} + x_z' \underbrace{(-\theta\, \delta_c)}_{\beta_c^{\tau}} + s_z \left( - \theta \rho_{ct} \right).$$

The last term varies by country, year and task, so neither set of fixed effects absorbs it. Adding and subtracting $\theta s_z \rho_{US,t}$, the piece $-\theta s_z \rho_{US,t}$ varies only by task and year and moves into $\alpha_{izt}$, leaving $s_z \beta_{ct}^{s}$ with $\beta_{ct}^{s} = -\theta(\rho_{ct} - \rho_{US,t})$ and $\beta_{US,t}^{s} = 0$. This is Equation (9). □

The step requires that LinkUp's coverage of a firm in a country does not differ across the firm's tasks within a year (Section 3.1). Because $s_z$ is measured with error (Section 4.6), the estimated $\beta_d^s$ are biased, in general toward zero. The same mismeasured $s_z$ enters every destination's term, but in a nonlinear model with fixed effects that does not make the bias a common factor: it can depend on how each destination's jobs are spread across titles, so the ordering of the estimates need not be that of the true coefficients. I do not derive the bias. The evidence on it is empirical: dropping the titles whose measure is least reliable (employer templates, clearance ladders, titles with few U.S. postings) moves most coefficients by a few hundredths, the Philippines' by up to 0.13, and leaves the Philippines lowest and Mexico next (Section 6.3).

Conditional on the total $\sum_c V_{iczt}$, the observed counts are multinomial with shares $\eta_{ict}\pi_{izct} / \sum_{c'} \eta_{ic't}\pi_{izc't}$: the placement shares of Lemma 1, reweighted by coverage. These equal Lemma 1's shares only if coverage is the same in every country. Because $\log \eta_{ict}$ varies only by firm, country and year, the reweighted shares have the same logit form with $\log \eta_{ict}$ added to $\gamma_{ict}$, so the slope coefficients are unchanged. A Poisson likelihood with a fixed effect for each firm × task × year has the same first-order conditions for the slope coefficients as this multinomial likelihood, so Poisson pseudo-maximum likelihood with these fixed effects estimates $\beta_c^{\tau}$ and $\beta_c^s$ consistently, and it remains consistent if the counts are not Poisson, provided the conditional mean is correct (Santos Silva and Tenreyro 2006). Firm–task–year cells in which the firm posts in one country only contribute zeros for the other countries, as the model implies they should. In words, the firm × task × year effect $\alpha_{izt}$ is how many jobs of the task the firm needs ($N_{izt}$, the multilateral term $\Phi_{izt}$ and the labor requirement), and the firm × country × year effect $\gamma_{ict}$ is the country's general attractiveness to the firm: wages, the general cost of working from that country, the firm's own premium and LinkUp's coverage.

## Migration

*Proof of Proposition 6.* With $H_d = (1-m_d)\bar{H}_d$ and $H_{US} = \bar{H}_{US} + \sum_{d'} m_{d'}\bar{H}_{d'}$, Equation (22) gives $\beta^s_d = (\theta/\sigma)[\log(1-m_d) + \log\bar{H}_d - \log L_d - \log H_{US} + \log L_{US}] + \text{const}$. Differentiating,

$$\frac{\partial \beta^s_d}{\partial m_d} = -\frac{\theta}{\sigma}\Big[\frac{1}{1-m_d} + \frac{\bar{H}_d}{H_{US}}\Big] < 0.$$

The first term is the loss of graduates in the destination and the second the gain in the United States. Emigration from $d$ also lowers $\beta^s_{d'}$ for every other destination $d'$, through the second term alone, by $(\theta/\sigma)\bar{H}_d/H_{US}$. □

## The Building–Application gap

For a Building task $z_B$ and an Application task $z_A$, applying Equation (8) to each and differencing gives

$$\log\frac{\pi_{iz_Bdt}}{\pi_{iz_B,US,t}} - \log\frac{\pi_{iz_Adt}}{\pi_{iz_A,US,t}} = -\theta\,(x_{z_B} - x_{z_A})'\delta_d \;-\; \theta\,(s_{z_B} - s_{z_A})(\rho_{dt} - \rho_{US,t}) \;-\; \theta\,\delta_{d,b,t}, \quad (23)$$

where $\delta_{d,b,t}$ is the Building-specific location cost that the measured attributes miss, added to $\log\tau_d(z_B)$.

*Proof of Proposition 3.* The terms $\omega_{id}$ and $\kappa_d$ are the same for both tasks and cancel. The right-hand side of Equation (23) is $\mathcal{L}_d + \mathcal{S}_d$. For (i), with $B_d = B_{US}$, Equation (22) gives $-\theta(\rho_{dt} - \rho_{US,t}) = (\theta/\sigma)[\log(H_d/L_d) - \log(H_{US}/L_{US})]$, so

$$\mathcal{S}_d = \frac{\theta}{\sigma}\,(s_{z_B} - s_{z_A})\Big[\log\frac{H_d}{L_d} - \log\frac{H_{US}}{L_{US}}\Big].$$

With $s_{z_B} > s_{z_A}$ this is negative if and only if $H_d/L_d < H_{US}/L_{US}$, increasing in $H_d/L_d$, and zero at equality. The location component $\mathcal{L}_d$ does not involve $H_d$, $L_d$ or $B_d$, so graduate supply places no restriction on it. If $\mathcal{L}_d = \mathcal{L}$ for every $d$, then $\Delta_d - \Delta_{d'} = \mathcal{S}_d - \mathcal{S}_{d'}$ and the gaps are ordered as the skill components are. Otherwise the ordering of $\Delta_d$ is that of $\mathcal{L}_d + \mathcal{S}_d$, and at $H_d/L_d = H_{US}/L_{US}$ the gap is $\mathcal{L}_d$, which need not be zero. For (ii), when $s_z\beta^s_d$ is included in the regression it absorbs $\mathcal{S}_d$, and the destination × Building coefficient estimates $\mathcal{L}_d$ only. For (iii), with time subscripts, the right-hand side of Equation (23) changes only through $\delta_{d,b,t}$ or $\rho_{dt} - \rho_{US,t}$; $\kappa_{dt}$ does not appear. □

## The productivity effect

*Proof of Proposition 5.* The cost of job $j$ in country $c$ satisfies $\Pr(C_{ijzc} > x) = \exp(-(x/k_{izc})^{\theta})$, independently across countries, so the lowest cost $C_{ijz}^{\min}$ satisfies $\Pr(C_{ijz}^{\min} > x) = \exp(-x^{\theta}\Phi_{iz})$. This is a Weibull distribution with mean $\Gamma(1{+}1/\theta)\,\Phi_{iz}^{-1/\theta}$. From Lemma 1, $\Phi_{iz} = k_{iz,US}^{-\theta}/\pi_{iz,US}$, so the mean equals $\Gamma(1+1/\theta)\,k_{iz,US}\,\pi_{iz,US}^{1/\theta}$: the cost of doing the job in the United States, scaled down by the U.S. share. By Equation (1) with $\tau_{US}(z) = 1$, the U.S. cost of task $z$ is

$$k_{iz,US} = a(z)\,(w_{US}^{H})^{s_z}(w_{US}^{L})^{1-s_z}\,\phi_{i,US}.$$

Weighted by the cost shares $\lambda_z$, the wage terms sum to

$$\log w_g \equiv \sum_{z\in g} \lambda_z \big[s_z \log w_{US}^{H} + (1-s_z)\log w_{US}^{L}\big],$$

the group's U.S. wage index, which is simply the group's wage when its workers are paid alike. The proposition holds $\phi_{i,US}$ fixed; if U.S. office costs change, $-\Delta\log\phi_{i,US}$ is added to the right side of Equation (11), and $w_g$ is pay net of those costs. Reading $w_g$ as worker pay requires that the wage, not the office cost, moves. With Cobb–Douglas cost shares $\lambda_z$ and zero profit,

$$\log p_g = \sum_{z\in g} \lambda_z \Big[\log\Gamma(1+1/\theta) + \log a(z) + \log w_g + \log\phi_{i,US} + \tfrac{1}{\theta}\log\pi_{z,US}\Big].$$

Differencing over time with $a(z)$ and $\phi_{i,US}$ fixed and using $\sum_z \lambda_z = 1$ gives Equation (11). □

The same formula answers two further questions. Measured against *task autarky*, every job in the United States ($\pi_{z,US} = 1$), the gain at fixed prices is $-\frac{1}{\theta}\sum_z \lambda_z \log\pi_{z,US}$. And for any change in location costs $\hat{\tau}_c(z)$, the observed shares and $\theta$ alone predict the new U.S. share, $\hat{\pi}_{z,US} = \big[\sum_c \pi_{zc}\hat{\tau}_c(z)^{-\theta}\big]^{-1}$, and with it the wage change.

*Proof of Proposition 7.* By the argument for Proposition 5, with a common wage $w_W$ the unit cost of task group $g$ satisfies $\log c_g = \sum_{z\in g}\lambda_z\big[\log\Gamma(1+1/\theta) + \log a(z) + \frac{1}{\theta}\log\pi_{z,US}\big] + \log w_W + \log\phi_{i,US}$, with $\phi_{i,US}$ held fixed as in Proposition 5. With Cobb–Douglas shares $\mu_g$ and zero profit in white-collar services, $\log p = \sum_g \mu_g \log c_g + \text{const}$. Differencing with $a(z)$ fixed and using $\sum_g \mu_g = \sum_{z\in g}\lambda_z = 1$ gives Equation (12). Part (i): with $\Delta\log p = 0$, every term $-\Delta\log\pi_{z,US} \geq 0$ and at least one is positive, so $\Delta\log w_W > 0$, and $w_W$ is the wage in every group. Part (ii) follows because there is one wage. Part (iii) rearranges Equation (12). The U.S. shares $\pi_{z,US}$ are themselves functions of $w_W$; the equation holds for their equilibrium values, which is what the data record. □

## The two channels for generative AI

*Proof of Proposition 4.* Junior hiring in country $c$ is $J_c = \sum_z a(z)\, j_z\, N_z \pi_{zc}$, summing over firms implicitly.

(i) From Lemma 1, $\partial \log \pi_{zd} / \partial \log \tau_d(z) = -\theta(1 - \pi_{zd}) < 0$ and $\partial \log \pi_{z,US} / \partial \log \tau_d(z) = \theta \pi_{zd} > 0$. A fall in $\tau_d(z)$ for every $z$ therefore raises $\pi_{zd}$ and lowers $\pi_{z,US}$ for every task, so $J_d$ rises and $J_{US}$ falls, and junior hiring in $d$ rises relative to the United States.

(ii) If AI lowers the junior labor requirement of task $z$ by a factor $1 - \psi_z$ in every country, it multiplies $k_{izc}$ by the same factor in every $c$ (to first order, by $1 - \psi_z j_z$), which cancels from $\pi_{zc}$ as $a(z)$ does in Lemma 1. Then $J'_c = \sum_z (1 - \psi_z) a(z) j_z N_z \pi_{zc}$ and, to first order, $\Delta \log J_c \approx -\sum_z \nu_{zc} \psi_z$ with $\nu_{zc} = a(z) j_z N_z \pi_{zc} / J_c$. Senior hiring $S_c$ is unchanged by construction, so the same expression gives the change in the log ratio of junior to senior hiring, $\Delta \log(J_c / S_c) \approx -\sum_z \nu_{zc} \psi_z$, which is what the regressions of Appendix D estimate. The junior share $q_c = J_c / (J_c + S_c)$ changes by less, $\Delta \log q_c \approx (1 - q_c)\, \Delta \log J_c$. □

Neutrality in (ii) needs the saving to leave the task's graduate share $s_z$ unchanged. If AI saves mostly non-graduate junior labor, $s_z$ rises and the skill term of Equation (8) moves graduate-intensive tasks slightly toward graduate-abundant countries; jobs can then move, but only through that term and only in that direction.

## The shape of the ladder

In each country a mid-level engineer is matched with $\ell_c$ juniors who execute codified tasks. The mid-level engineer draws problems and solves those in the set of knowledge $k_c$ within reach, a share $F(k_c)$. The rest are escalated, and each takes a senior $h$ units of time, so one mid-level engineer generates demand for $q_c = h[1 - F(k_c)]$ seniors (Garicano and Rossi-Hansberg 2006). This gives Equation (13). Write $D_c = \ell_c + 1 + q_c$, so the shares of the three rungs are $\ell_c / D_c$, $1 / D_c$ and $q_c / D_c$.

*Proof of Proposition 8.* Let $d \log \ell_c = -\hat{\ell}_c \leq 0$ and $d \log q_c = -\hat{q}_c \leq 0$, not both zero. Then $d \log D_c = -(\ell_c \hat{\ell}_c + q_c \hat{q}_c) / D_c < 0$, so the mid-level share $1 / D_c$ rises, which is (i). The junior share changes by

$$d \log \frac{\ell_c}{D_c} = -\hat{\ell}_c + \frac{\ell_c \hat{\ell}_c + q_c \hat{q}_c}{D_c} = \frac{-\hat{\ell}_c (1 + q_c) + q_c \hat{q}_c}{D_c},$$

which is negative if and only if $\hat{q}_c / \hat{\ell}_c < (1 + q_c) / q_c$. The senior share changes by

$$d \log \frac{q_c}{D_c} = -\hat{q}_c + \frac{\ell_c \hat{\ell}_c + q_c \hat{q}_c}{D_c} = \frac{-\hat{q}_c (1 + \ell_c) + \ell_c \hat{\ell}_c}{D_c},$$

which is negative if and only if $\hat{q}_c/\hat{\ell}_c > \ell_c/(1+\ell_c)$. With $\hat{q}_c = 0$ the junior share falls and the senior share rises, which is (ii); both shares fall exactly when both inequalities hold, which is (iii). □

In a team that spans countries (Antràs et al. 2006), problems that need tied or doctoral knowledge are escalated to U.S. seniors. Seniors outside the United States therefore handle the problems that are hard for the local team but not frontier problems, and these are the problems that a tool giving access to codified knowledge is most likely to put within reach: $F'(k_d)$ is large over the relevant range and $\hat{q}_d$ is large. U.S. seniors handle the frontier and doctoral problems, for which $\hat{q}_{US}$ is small, so the U.S. ladder is in case (ii). This is an assumption about where AI's knowledge is useful, not a result; the data in Table 15 are consistent with it.

Proposition 8 compares steady states. In the dynamic model of Friebel et al. (2026), juniors become seniors with some probability, so cutting junior hiring today lowers the supply of seniors later. A productivity shock then leads to a short-run freeze in junior hiring, after which the firm restores its long-run span, possibly after oscillating; a shock to the speed of learning changes the long-run span. A fall in $q_c$ plays the role of their learning shock, since it lets mid-level engineers resolve what used to require seniority. I use their dynamic result as given and do not derive it here.

Online Appendix

Splitting the Stack: Generative AI and the Cross-Country Division of Technical Work

Hitanshu Pandit

# D Junior and senior software hiring after ChatGPT

This appendix gives the junior-hiring analysis summarized in Section 8.3: the design of Westby et al. (2026) applied separately to the firms' U.S. and Indian offices, the event study, the shape of the experience ladder, the shift-share decomposition, and the pre-specified variants.

## D.1 The design

The second margin is hiring by experience. Take the junior software-developer jobs a firm posts in the United States and compare them with its senior software-developer jobs, before and after November 2022. If juniors fell relative to seniors, that might be AI, or it might be higher interest rates, which slowed junior hiring everywhere. So compare the same junior-versus-senior change in the firm's other technical occupations, which faced the same interest rates but whose tasks overlap less with what language models do. The difference between the two changes is the estimate. This is the design of Westby et al. (2026) for U.S. vacancies; I apply it separately to each firm's U.S. and Indian offices.

Postings that state a minimum experience requirement are split into the three rungs fixed in advance in Section 3. For firm $f$, country $c$, occupation group $g$ (software developers, O*NET 15-1252, or all other technical occupations), rung $l$ and month $t$, I estimate by Poisson pseudo-maximum likelihood

$$\mathbb{E}[n_{fcglt}] = \exp\big(\alpha_{fcgl} + \mu_{fcgt} + \lambda_{clt} + \beta_J \,\mathrm{Junior}_l \times \mathrm{Post}_t \times \mathrm{SW}_g + \beta_M \,\mathrm{Mid}_l \times \mathrm{Post}_t \times \mathrm{SW}_g\big), \quad (24)$$

separately by country. The first set of fixed effects holds fixed each firm's usual mix of rungs in each occupation group, the second absorbs everything that moves the firm's hiring in that group in a given month, and the third absorbs anything common to a rung in the country and month. A negative $\beta_J$ means that after ChatGPT the firm's junior software postings fell relative to its senior ones by more than in its other technical work. The window is Westby et al.'s, November 2021 to October 2023, and standard errors are clustered by firm. The specification is one of ten pre-specified variants that differ in how exposure to AI is measured

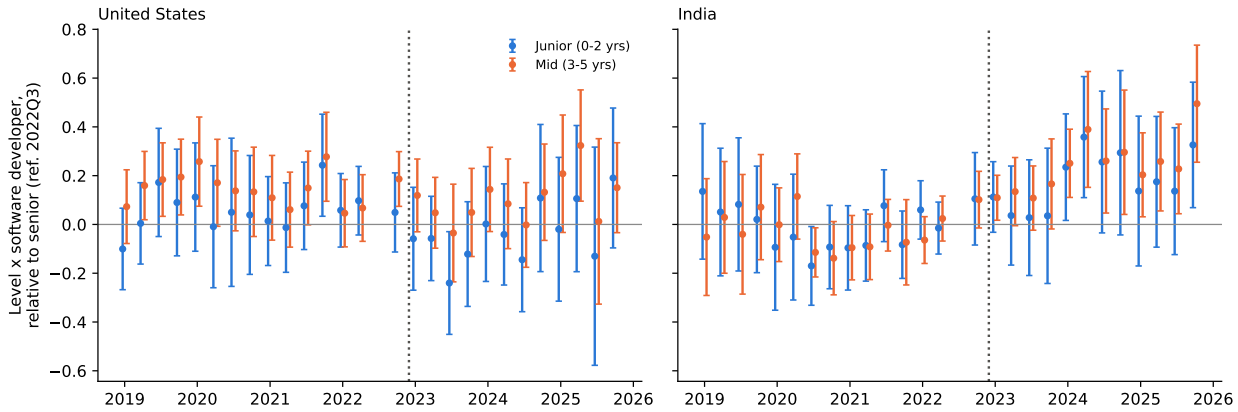



Figure 18: Junior (0–2 years) and mid-level (3–5 years) software-developer postings relative to senior postings and to other technical occupations, by quarter, U.S. and Indian operations of the same firms (reference 2022Q3). Coefficients from the event-study version of Equation (24), 2019–2025, with 95 percent intervals; standard errors clustered by firm. The dotted line marks the release of ChatGPT.

and which occupations serve as the comparison (Appendix Table 16); the software-developer comparison is the only one with the expected sign that survives the adjustment for testing ten.

## D.2 Results

In the firms' U.S. offices, junior software-developer postings fell by 0.215 log points (s.e. 0.063) relative to senior postings and to junior postings in other technical occupations, a decline of 19 percent (Table 14). Mid-level postings did not move. With Westby et al.'s own two-rung split, the estimate is −0.155 (s.e. 0.052), close to their −0.136 for U.S. vacancies as a whole. In the same firms' Indian offices there is no such decline: −0.021 (s.e. 0.073) for juniors. Before the release the junior coefficients are flat in both countries, and a false release date a year earlier gives zero.

The U.S. decline was short (Figure 18). Junior postings were close to their usual level for two quarters after the release, fell by 0.24 (s.e. 0.11) in the third quarter of 2023, and were back around zero from 2024. Part of the pooled estimate also comes from unusually strong junior hiring at the end of 2021, inside the comparison window. Averaged over the year after the release, the decline is −0.08, with a 95 percent interval from −0.23 to +0.08. What the data support is a dip in U.S. entry-level software postings in 2023, not a lasting fall. Friebel et al. (2026) predict exactly this when AI raises productivity: a short freeze in junior hiring, after which firms rebuild the bottom of the ladder because today's juniors are the seniors they will need.

Table 14: Junior and mid-level software-developer postings after ChatGPT, U.S. and Indian operations of the same firms

| | United States | India | India − U.S. |
|---|---|---|---|
| *A. Westby et al. window, November 2021 – October 2023* | | | |
| Junior (0–2 years) × Post × SW | −0.215*** | −0.021 | 0.194** |
| | (0.063) | (0.073) | (0.087) |
| Mid (3–5 years) × Post × SW | −0.043 | 0.067 | 0.109 |
| | (0.060) | (0.053) | (0.074) |
| Holm-adjusted $p$, junior | 0.002 | 0.779 | 0.052 |
| *B. Placebo: false date November 2021, window November 2020 – October 2022* | | | |
| Junior × Post × SW | 0.010 | 0.009 | −0.001 |
| | (0.078) | (0.077) | (0.127) |
| *C. Pre-trends, 2019Q1–2022Q2 (joint Wald test, p-value)* | | | |
| Junior × SW | 0.28 | 0.12 | |
| Mid × SW | 0.06 | $< 0.001$ | |
| *D. Two-level split (junior $< 4$ years), same window* | | | |
| Junior × Post × SW | −0.155*** | 0.006 | 0.160*** |
| | (0.052) | (0.054) | (0.061) |
| Firms (panel A) | 3,190 | 2,155 | 3,452 |

*Notes:* PPML of Equation (24) on firm × occupation group × level × month counts of technical postings with a stated minimum experience (LinkUp; KPMG, Deloitte and Infosys excluded). SW: software developers (O*NET 15-1252); comparison: all other technical occupations. Senior (6+ years) is the reference level. Fixed effects firm × group × level, firm × group × month and level × month (by country); India − U.S. pools the arms with all fixed effects interacted with country. Standard errors clustered by firm in parentheses; stars use unadjusted $p$-values (Holm-adjusted $p$ reported separately): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel C: quarterly event study, reference 2022Q3. Panel D: S1 of the pre-specified family (Appendix Table 16).

India changed in a different way. Between 2022 and 2024 the mid-level share of Indian software-developer postings rose from 46 to 65 percent, while both the junior and the senior shares fell (Table 15); the new shape held through 2025. Other Indian technical occupations moved the same way, by less, and in the United States the middle barely moved. In the language of the model this is a diamond rather than a pyramid (Equation (13)), in the Indian office and not in the U.S. one. It should not be read as an effect of ChatGPT: the shift came in 2024, India's mid-level share was already rising before 2022, and it coincides with a slowdown in hiring of new graduates across Indian technical work.

Table 15: The shape of the ladder: shares of software-developer and other technical postings at each rung, U.S. and Indian operations

| | | Software developers | | | Other technical occupations | | |
|---|---|---|---|---|---|---|---|
| | Year | Junior | Mid | Senior | Junior | Mid | Senior |
| United States | 2019 | 30.8 | 47.6 | 21.6 | 31.8 | 45.7 | 22.4 |
| | 2022 | 39.5 | 41.8 | 18.6 | 32.2 | 47.2 | 20.6 |
| | 2023 | 33.3 | 45.1 | 21.6 | 32.3 | 46.3 | 21.3 |
| | 2024 | 38.3 | 42.4 | 19.3 | 32.5 | 46.0 | 21.5 |
| | 2025 | 35.8 | 44.2 | 20.0 | 30.2 | 45.7 | 24.0 |
| India | 2019 | 25.5 | 51.8 | 22.8 | 28.6 | 49.0 | 22.5 |
| | 2022 | 27.7 | 46.1 | 26.2 | 29.2 | 44.0 | 26.8 |
| | 2023 | 19.4 | 53.9 | 26.7 | 25.7 | 46.9 | 27.5 |
| | 2024 | 18.1 | 64.9 | 17.0 | 22.5 | 55.1 | 22.4 |
| | 2025 | 17.4 | 64.5 | 18.0 | 22.0 | 54.2 | 23.8 |

*Notes:* Percent of technical postings with an O*NET code and a stated minimum experience requirement, by year (LinkUp; KPMG, Deloitte and Infosys excluded). Junior: 0–2 years; mid: 3–5; senior: 6 or more. Software developers: O*NET 15-1252. Rows sum to 100 within each occupation group.

Indian employers also ask for more experience for the same job. Comparing the same firm, title and year in the two countries across all technical titles, Indian postings require 0.87 more years on average (s.e. 0.21; for software developers see Appendix D.3); only 6 percent of Indian software-developer postings ask for one year, against 15 percent in the United States. A fixed two-year threshold therefore picks a slightly higher rung in India. Defining the rungs instead by each country's own distribution of requirements leaves the results unchanged ($-0.214$ in the United States, $+0.026$ in India).

## D.3 What is different about junior postings in India

Indian software postings are less junior than U.S. ones, and after ChatGPT they became less junior still. India could get a different mix of work, more of the kinds of jobs that are

rarely junior in any country, or the same kind of job could simply ask for more experience in India. The shift-share analysis of Westby et al. (2026) separates the two. The junior share of a country's postings is an average over kinds of job, weighted by how common each kind is, so a difference between two countries, or a change over time, splits into a part that comes from the mix and a part that comes from the junior share within the same kind of job. I use the seven task categories as the kinds of job, and the firm as a check (Table 11).

The answer is the requirement, not the mix. In the year before ChatGPT (November 2021 to October 2022), 27.8 percent of Indian software-developer postings were junior against 39.8 percent in the United States, a gap of 11.9 points. Indian software work is far more concentrated in technical application, two-thirds of it against half in the United States, but that difference in mix accounts for less than one point of the gap; the rest is the same task asking for more experience in India. Holding the employer and the title fixed sharpens the comparison: in the same firm, for the same job title, Indian postings ask for 1.1 more years of experience and are 15 points less likely to be junior. Part of the gap does come from which firms post where, because the firms that post most in India also ask for more experience everywhere (Panel A, last column).

The changes after ChatGPT are of the same kind. In the first year the junior share of software postings fell by 6.4 points in the United States and 8.1 points in India, and in both countries almost all of the fall happened within task categories; firms did not move toward kinds of work that are rarely junior. The two countries then went separate ways. By 2024–25 the U.S. fall had mostly reversed, to 2.8 points below the level before the release. India's had deepened to 10.1 points, and it was not confined to software: the junior share of India's other technical postings fell from 29.5 to 22.2 percent over the same years, while in the United States it barely moved. That is why the comparison with other occupations in Table 14 finds nothing specific to Indian software developers: the decline in India is broad, and it is a decline in how junior the same work is allowed to be.

## D.4 Supporting evidence and what fails

Some of the pre-specified variants point the same way without surviving the adjustment for testing ten. Using non-software engineers as the comparison gives −0.147 (s.e. 0.072). Occupations whose tasks people more often hand to Claude, and more often for automation than for help, show larger junior declines (−0.053 and −0.057 per standard deviation of the Anthropic Economic Index measures). Measures of what models could do in principle, such as the GPT-4 exposure ratings of Eloundou et al. (2024), show no gradient. The decline sits where generative AI was used first and most, in writing code.

Other designs, also written down in advance, neither confirm nor overturn this (Appendix P). Dating each task category by when a language model could first do its core work gives imprecise estimates. The full-text categories give U.S. junior declines in model building and technical application that do not survive the multiple-testing adjustment. Instruments for how much work a firm places in India and for AI use are too weak to use.

One prediction fails outright. If the U.S. dip reflects codified junior work that had not yet moved to India, a firm that had already moved more of its application work to India should have had less of it left in the United States, and a smaller U.S. junior decline. It did not: the interaction with the firm's pre-2022 Indian share has the wrong sign ($-0.086$, s.e. 0.083), and in the firms that had placed most work in India the U.S. declines are larger, not smaller, a trend that began before the release.

### D.5 Why the two countries differ

The model offers one reading of the contrast, an interpretation rather than a separately identified channel. By Proposition 4(ii), the codified slice of every occupation has already moved to India, so a comparison of software with other occupations has little to detect there, while in the United States software stands out. The shift-share evidence fits: India's junior decline is broad and within tasks (Appendix D.3). The reading explains the contrast between the two countries, not differences among firms, where the test fails.

## E The levels design: the pre-specified family and the instruments tried

Table 16 reports the ten designs specified in a pre-analysis plan written before estimation (available from the author), for the U.S. arm, estimated on the same cells (firm $\times$ country $\times$ occupation $\times$ level $\times$ month, two levels split at four years) and fixed effects, with Holm-adjusted p-values across the ten. The pre-specified rule required a Holm-adjusted $p < 0.05$ and a placebo with $p \geq 0.10$; it did not state a sign. S3 meets it with a positive coefficient: human raters place software developers near the bottom of technical exposure, so S3 restates S1 on an inverted scale. The Anthropic Economic Index measures (S5–S7) use task statements coded in O*NET-SOC 2010, mapped to 2019 codes with the official crosswalk. No member is significant in the Indian arm.

Table 16: Pre-specified family of exposure designs, U.S. arm, Westby et al. window

| Design | Estimate | Holm $p$ | Placebo | Meets rule |
|---|---|---|---|---|
| S1 Software developers vs other technical | $-0.155^{***}$ | 0.025 | 0.030 | Yes |
| | (0.052) | | (0.053) | |
| S2 GPT-4 exposure ($\beta$), per s.d. | $-0.035$ | 0.538 | $-0.019$ | No |
| | (0.026) | | (0.031) | |
| S3 Human exposure ($\beta$), per s.d. | $0.067^{***}$ | $< 0.001$ | $-0.005$ | Yes (wrong sign) |
| | (0.016) | | (0.023) | |
| S4 GPT-4 direct exposure ($\alpha$), per s.d. | $-0.044^{*}$ | 0.347 | $-0.016$ | No |
| | (0.026) | | (0.028) | |
| S5 AEI usage intensity, per s.d. | $-0.053^{**}$ | 0.227 | $-0.010$ | No |
| | (0.025) | | (0.029) | |
| S6 AEI automation share, per s.d. | $-0.057^{**}$ | 0.145 | $-0.004$ | No |
| | (0.025) | | (0.023) | |
| S7 AEI automation, with usage | $-0.048$ | 0.538 | 0.029 | No |
| | (0.036) | | (0.051) | |
| S8 Software developers vs engineers | $-0.147^{**}$ | 0.227 | 0.031 | No |
| | (0.072) | | (0.071) | |
| S9 Software developers vs hardware/network | $-0.249^{***}$ | 0.026 | $0.208^{**}$ | No (placebo) |
| | (0.084) | | (0.106) | |
| S10 Firm-level GPT-4 exposure | $-0.013$ | 0.538 | 0.009 | No |
| | (0.010) | | (0.008) | |

*Notes:* PPML, Junior ($< 4$ years) $\times$ Post $\times$ exposure, fixed effects firm $\times$ occupation $\times$ level, firm $\times$ occupation $\times$ month and level $\times$ month; S10 is a firm-level weighted regression of the change in the junior share on firm exposure. Placebo: false date November 2021. Standard errors clustered by firm in parentheses; stars use unadjusted $p$-values: $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table 17 summarizes two instruments, each specified before estimation and reported whatever it showed. The first instruments the change in India's share of the firm's technical postings, with the change in the U.S. junior share as the outcome (one observation per firm); the second instruments whether a posting names a generative-AI tool. Neither has a first-stage F statistic above 3.3.

Table 17: Instruments tried for the levels design

| Instrument | Endogenous variable | First-stage F | Units |
|---|---|---|---|
| Shift-share: pre-period occupation shares $\times$ leave-one-out change in India's share of LinkUp postings by occupation | Change in India share of the firm's postings | 0.7 | 1,062 firms |
| Public-code abundance (The Stack) of the programming languages a posting names $\times$ Post | Posting names a generative-AI tool | 3.0 (U.S.), 3.3 (India) | 2.9 million postings |

*Notes:* Robust (firm-level designs) or firm-clustered (posting-level design) first-stage F statistics. The language instrument's reduced form fails its placebo and pre-trend test. Each instrument was specified in a pre-analysis plan before estimation.

# F Junior hiring in Building and Application versus general tech and engineering work

India's information technology and services sector reduced fresher and junior hiring in 2023 and 2024. Such a contraction would lower junior hiring across all technical work in India, whatever its category, from about the time of ChatGPT's release. To separate a change specific to Building and Application work from a general one, I compare Building and Application jobs with general-tech and engineering jobs at the same firms: lookup *general tech* titles plus mechanical, civil, electrical, chemical, structural, manufacturing, industrial, process, aerospace, automotive, HVAC, piping and controls engineers. The comparison is made within firm, so that any shock to a firm's junior hiring in a country is differenced out, and it is estimated both with the two-level split of Westby et al. (2026) and with three levels (junior 0–2 years, mid 3–5, senior 6 or more), fixed before estimation.

Because a PPML specification with firm × group × level, firm × group × period and firm × level × period fixed effects leaves each firm's twelve cells nearly saturated and does not converge, I estimate the triple difference directly on shares. For each firm with at least five jobs in both the Building-and-Application group and the comparison group, before and after November 2022, I compute the change in the share of its Building and Application jobs at a given level minus the same change for its comparison jobs, and average across firms weighted by Building and Application jobs, with heteroskedasticity-robust standard errors. Firm-specific shocks to junior hiring cancel within firm.

Table 18: Change in the junior share of Building and Application jobs relative to general-tech and engineering jobs, same firm

| | United States | India | Mexico |
|---|---|---|---|
| *Two levels (junior < 4 years)* | | | |
| Application, junior | 1.15 | 2.14 | 5.56 |
| | (2.29) | (7.32) | (3.47) |
| Building, junior | −2.71 | 8.03 | — |
| | (4.25) | (8.36) | |
| *Three levels (junior 0–2, mid 3–5; senior 6+ reference)* | | | |
| Application, junior | 2.13 | 3.95 | 11.36* |
| | (1.65) | (6.66) | (6.88) |
| Application, mid | −2.79** | −2.39 | −10.09** |
| | (1.35) | (3.87) | (5.01) |
| Firms | 643 | 109 | 35 |

*Notes:* Percentage points; robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each entry is the firm-weighted average of (change in the share of the firm's Building and Application jobs at the level, pre to post November 2022) minus (the same change for its general-tech and engineering jobs). Firms need at least five jobs in each group and period. The United Kingdom and the Philippines have too few qualifying firms (18 and 15) to report; Building cells outside the United States and India are too small.

None of the Indian estimates is distinguishable from zero, and the point estimates for junior Application jobs are positive: the junior share of Building and Application jobs did not fall by more than that of comparable general-tech and engineering jobs at the same firms. The comparison rests on 109 Indian firms and is imprecise, so it neither confirms nor rules out a component specific to Building and Application work; it does not support attributing the Indian decline to that work.

# G Data construction

Panel firms are mapped from `unique_firm_id` to LinkUp company identifiers through the firm lookup used throughout. The excluded firms are `firm_002877` (KPMG), `firm_003084` (Deloitte & Touche) and `firm_007589` (Infosys), identified by a whole-word name match checked by hand.

Table 19: Sample funnel: from the firm universe to the principal estimate

| Step | Firms | Postings or jobs |
|---|---|---|
| Firms in LinkUp's U.S.–India technical-posting frame, 2019–2025 (technical postings) | 4,552 | 13,577,769 |
| of which advertise technical jobs in both countries in 2019–2025 | 3,558 | |
| Excluding KPMG, Deloitte and Infosys | 4,549 | |
| no technical posting with a full description (never labelled) | −33 | |
| labelled postings, none matched to a title and creation date | −14 | |
| Principal sample: English distinct jobs, five countries | 4,502 | 6,089,734 |
| of which observed only through British, Mexican or Philippine postings | 23 | |
| Firm × country × category × year cells | | 762,265 |
| remaining after Poisson drops all-zero fixed-effect groups | 4,502 | 282,703 |
| firms with a year of jobs in both the United States and a destination | 3,692 | |
| *Robustness:* including KPMG, Deloitte and Infosys (U.S.–India jobs only) | 4,505 | 6,196,796 |
| Building − Application, India / Philippines / Mexico / United Kingdom | −0.198 / −0.463 / −0.296 / −0.012 | |
| principal estimate | −0.199 / −0.464 / −0.296 / −0.013 | |

*Notes:* The counts rebuild the principal sample of Table 5 and reproduce its jobs, firms and cells exactly. Firms without a full description have a median of 2 technical postings (maximum 288); firms lost at the title and date match have a median of 114 (maximum 23,145). The three excluded firms' British, Mexican and Philippine postings were not labelled, so in the robustness row they enter only the U.S.–India comparison; standard errors 0.040, 0.051, 0.044 and 0.029 in both rows. In LinkUp's multicountry extract of all postings, 4,534 of 4,544 firms advertise in both countries; the rule that selected the 4,552-firm frame is not documented in the replication files.

One pass over LinkUp's 291 raw `job_record` shards keeps every posting of panel firms in the five countries. Duplicate `JOB_HASH`es across shards are removed, keeping the earliest record. A distinct job is a unique (firm, country, title, creation date); titles are normalized by lower-casing and replacing every non-alphanumeric character with a space. Location copies differ by country and task: a U.S. Building job appears 2.84 times on average against 2.59 for a U.S. Application job, and an Indian Building job 1.66 times against 1.98 (jobs in lookup titles). Raw posting counts would therefore overstate the U.S. share of Building work.

The location-cost terms are applied case-insensitively to U.S. descriptions. *Clearance:* a security, secret, government, federal or DoD clearance; an active, current, obtained, maintained or held (security) clearance; "clearance required"; top secret; TS/SCI; U.S. citizen or citizenship; ITAR; public trust. *On-site:* on-site, onsite, in-office, in person. *Client-facing:* client-facing, customer-facing, working directly with (our) clients or customers, client meetings or engagements, stakeholder management. *Travel:* travel. *Remote:* remote. A bare "clearance" pattern is not used because it matches phrases such as "clearance of backlog". The exact expressions are in the replication package.

The skill-intensity terms are applied case-insensitively to U.S. descriptions of titles in the six-category and ordinary technical lookups. *PhD:* PhD, Ph.D., doctorate, doctoral degree. *Master's:* master's, masters degree, master of science, engineering, arts or computer, M.S. degree or M.S. in, MS degree, MS in a technical field, MSc, M.Tech ("MS" alone was dropped because it matches MS Office). *Research record:* publications in, publication record, published research or papers, peer-reviewed, first-author, top-tier journals, conferences or venues (a bare "publication" was dropped because it matched "DoD Publication 8570").

Each category's seed vocabulary (listed in the replication package) is matched, token by token, against the twenty most distinctive terms of each group; a group needs at least one match. Abbreviated: *Model building:* PyTorch, TensorFlow, deep learning, neural networks, transformer architecture, LLM training, fine-tuning, pre-training, foundation model, distributed training, CUDA, machine learning engineer, research or applied scientist. *Data and ML infrastructure:* MLOps, Kubeflow, MLflow, model registry, model serving and monitoring, feature store, CI/CD and data pipelines, ETL, Airflow, Spark, Kafka, Kubernetes, Docker, Terraform, platform, site-reliability and DevOps engineer. *Technical application:* OpenAI and Anthropic APIs, LangChain, LlamaIndex, prompt engineering, retrieval-augmented generation, vector databases, semantic search, function calling, agentic workflows, API integration. *Functional application:* AI strategy, adoption, transformation and roadmap, digital transformation, AI product or program manager, change management, business case, stakeholder management, AI consultant. *Product and testing:* conversational and interaction design, human–AI interaction, UX research, usability testing, human-centered design. *Governance and security:* AI governance, responsible AI, AI ethics and risk management, model risk management, algorithmic fairness, AI safety, regulatory compliance, data privacy, GDPR, model validation, privacy engineer.

Table 20: Distinct focal-category jobs by year and country

| Year | United States | India | United Kingdom | Mexico | Philippines | India share (%) |
|---|---|---|---|---|---|---|
| 2019 | 102,588 | 18,197 | 10,046 | 1,762 | 1,095 | 15.1 |
| 2020 | 80,781 | 16,913 | 6,287 | 1,400 | 811 | 17.3 |
| 2021 | 144,535 | 38,267 | 13,950 | 4,072 | 2,155 | 20.9 |
| 2022 | 151,940 | 45,182 | 16,763 | 4,293 | 2,832 | 22.9 |
| 2023 | 99,817 | 33,046 | 14,416 | 3,816 | 2,815 | 24.9 |
| 2024 | 81,456 | 35,662 | 10,951 | 3,504 | 2,900 | 30.4 |
| 2025 | 71,056 | 38,342 | 8,469 | 3,362 | 2,828 | 35.0 |
| Total | 732,173 | 225,609 | 80,882 | 22,209 | 15,436 | 23.6 |

*Notes:* Distinct jobs in the six focal categories of the fixed title lookup. India share: India's percent of combined U.S. and Indian jobs. Year-to-year changes partly reflect LinkUp's scraping coverage.

Table 21: Distinct jobs by task category and country, 2019–2025

| | Titles | United States | India | United Kingdom | Mexico | Philippines |
|---|---|---|---|---|---|---|
| *Building* | | | | | | |
| T1 Model building | 1,035 | 27,366 (3.7) | 6,593 (2.9) | 2,314 (2.9) | 428 (1.9) | 444 (2.9) |
| T2 Data and ML infrastructure | 1,843 | 36,911 (5.0) | 12,455 (5.5) | 3,250 (4.0) | 943 (4.2) | 378 (2.4) |
| *Application* | | | | | | |
| T3 Technical application | 1,895 | 90,294 (12.3) | 20,634 (9.1) | 5,661 (7.0) | 1,405 (6.3) | 685 (4.4) |
| T5 Functional application | 12,922 | 479,198 (65.4) | 142,340 (63.1) | 59,248 (73.3) | 15,436 (69.5) | 12,141 (78.7) |
| T4 Product and testing | 1,461 | 63,550 (8.7) | 36,227 (16.1) | 6,635 (8.2) | 3,503 (15.8) | 1,129 (7.3) |
| T6 Governance and security | 1,588 | 34,854 (4.8) | 7,360 (3.3) | 3,774 (4.7) | 494 (2.2) | 659 (4.3) |
| *Benchmark* | | | | | | |
| T7 General tech | 3,116 | 54,033 | 11,164 | 3,599 | 1,766 | 1,796 |
| Focal categories, total | 20,744 | 732,173 | 225,609 | 80,882 | 22,209 | 15,436 |
| All raw postings | | 45,362,919 | 5,872,110 | 3,425,189 | 867,043 | 441,002 |

*Notes:* Distinct jobs (firm × country × title × creation date) in titles of the fixed lookup (Section H). In parentheses: percent of the country's focal-category jobs. "Titles" is the number of lookup titles in each category. "All raw postings" counts every unique LinkUp record of the sample firms in the country, in any title. KPMG, Deloitte and Infosys excluded.

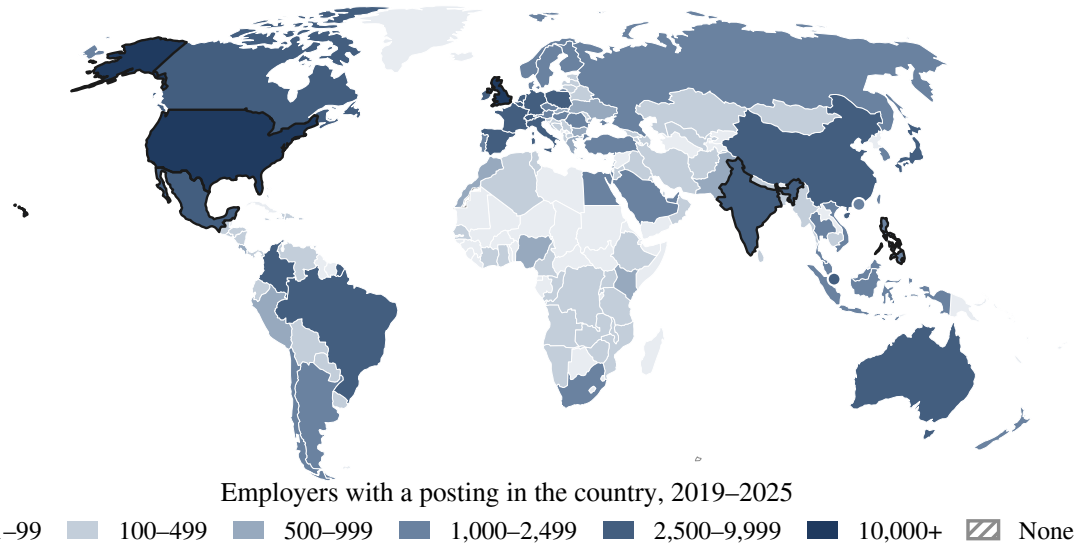



Figure 19: Employers in LinkUp with at least one job posting located in each country, 2019–2025. An employer is a LinkUp company identifier (one scraped career site), counted in every country in which it posts. Hatched: no posting. Outlined: the United States and the four destinations studied here. Singapore and Hong Kong, too small to draw at this scale, are shown as dots. Postings with no country (6.9 percent) are excluded. Source: LinkUp raw job records (291 shards, 349.5 million postings).

Table 22: Required experience by category and country, distinct jobs, 2019–2025

| Category | Country | Jobs stating (%) | Mean years | Median | Junior (%) | Senior (%) |
|---|---|---|---|---|---|---|
| T1 Model building | United States | 62 | 4.1 | 3 | 34.3 | 20.3 |
| | India | 64 | 4.3 | 3 | 26.7 | 20.5 |
| | United Kingdom | 22 | 3.6 | 3 | 33.3 | 11.4 |
| | Mexico | 52 | 3.9 | 3 | 29.6 | 12.6 |
| | Philippines | 48 | 3.3 | 3 | 37.4 | 8.4 |
| T2 Data and ML infrastructure | United States | 67 | 4.1 | 3 | 36.3 | 21.4 |
| | India | 67 | 4.6 | 4 | 22.1 | 25.2 |
| | United Kingdom | 23 | 4.3 | 3 | 28.4 | 18.9 |
| | Mexico | 53 | 4.1 | 4 | 22.9 | 15.1 |
| | Philippines | 46 | 3.3 | 3 | 45.7 | 10.9 |
| T3 Technical application | United States | 64 | 4.1 | 3 | 34.7 | 23.0 |
| | India | 69 | 4.5 | 4 | 24.6 | 27.4 |
| | United Kingdom | 18 | 4.1 | 4 | 27.0 | 14.9 |
| | Mexico | 57 | 3.3 | 3 | 34.2 | 9.4 |
| | Philippines | 53 | 3.1 | 3 | 45.0 | 8.8 |
| T5 Functional application | United States | 52 | 3.8 | 3 | 37.3 | 16.7 |
| | India | 54 | 4.2 | 4 | 31.1 | 22.7 |
| | United Kingdom | 18 | 4.0 | 3 | 32.6 | 15.3 |
| | Mexico | 42 | 3.5 | 3 | 37.7 | 11.1 |
| | Philippines | 49 | 3.4 | 3 | 43.7 | 9.9 |
| T4 Product and testing | United States | 66 | 3.5 | 3 | 41.8 | 14.7 |
| | India | 65 | 4.0 | 3 | 27.5 | 19.1 |
| | United Kingdom | 18 | 4.1 | 3 | 30.3 | 13.7 |
| | Mexico | 47 | 3.5 | 3 | 32.8 | 13.5 |
| | Philippines | 53 | 3.4 | 3 | 40.8 | 10.6 |
| T6 Governance and security | United States | 67 | 4.0 | 3 | 34.2 | 20.4 |
| | India | 62 | 4.1 | 3 | 31.1 | 19.0 |
| | United Kingdom | 19 | 4.3 | 3 | 27.1 | 15.2 |
| | Mexico | 46 | 3.5 | 3 | 35.1 | 9.3 |
| | Philippines | 55 | 3.0 | 3 | 42.4 | 5.8 |
| T7 General tech | United States | 55 | 2.8 | 2 | 58.4 | 7.3 |
| | India | 65 | 3.1 | 2 | 52.1 | 12.5 |
| | United Kingdom | 20 | 2.9 | 2 | 50.1 | 5.0 |
| | Mexico | 41 | 2.1 | 2 | 70.2 | 3.8 |
| | Philippines | 52 | 2.2 | 2 | 73.7 | 1.8 |

*Notes:* Distinct jobs (firm × country × title × creation date) in the fixed title lookup. "Jobs stating": share with a parsed minimum experience requirement; the remaining columns use those jobs. Junior: 0–2 years; senior: 6 or more. KPMG, Deloitte and Infosys excluded.

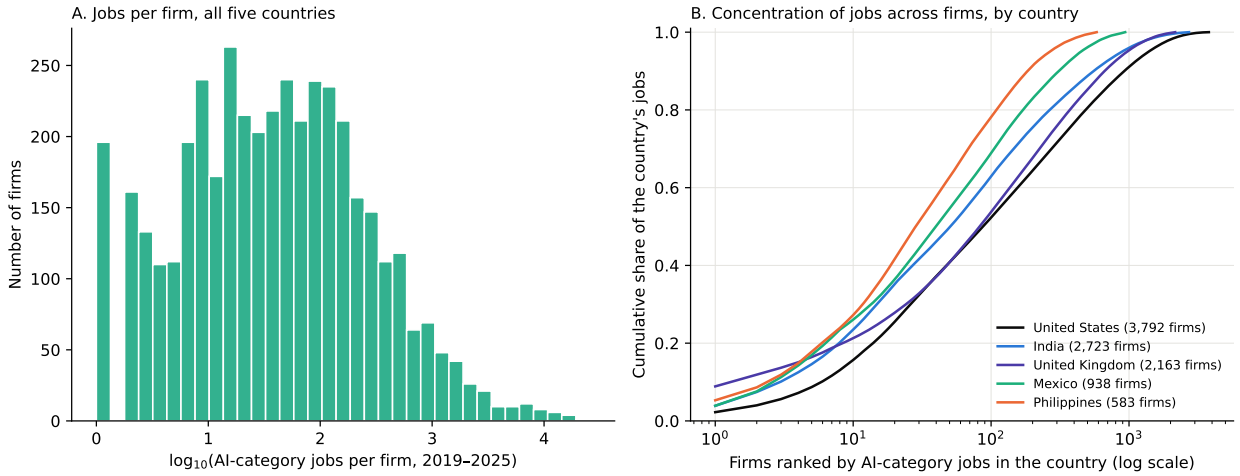



Figure 20: Distribution of focal-category jobs across firms. Panel A: distinct focal-category jobs per firm, 2019–2025, all five countries, log scale (4,210 firms). Panel B: cumulative share of each country's focal-category jobs held by its largest firms, firms ranked by jobs in that country.

Table 23: Required experience by occupation, United States and India, technical postings 2019–2025

| | United States | | | India | | |
|---|---|---|---|---|---|---|
| Occupation (O*NET) | Mean years | Junior (%) | Senior (%) | Mean years | Junior (%) | Senior (%) |
| Software Developers | 3.9 | 36.6 | 19.5 | 4.4 | 21.0 | 23.2 |
| Management Analysts | 4.0 | 34.8 | 19.8 | 4.2 | 31.3 | 22.8 |
| Computer and Information Systems Managers | 5.5 | 15.7 | 39.6 | 5.7 | 11.0 | 31.4 |
| Information Security Analysts | 4.3 | 31.2 | 22.1 | 4.5 | 22.1 | 24.2 |
| Computer User Support Specialists | 2.9 | 52.7 | 8.8 | 3.4 | 40.4 | 13.0 |
| Market Research Analysts and Marketing Specialists | 3.5 | 40.9 | 12.4 | 3.4 | 38.4 | 13.4 |
| Data Scientists | 3.9 | 37.2 | 17.3 | 4.1 | 30.5 | 20.4 |
| Information Technology Project Managers | 4.6 | 23.4 | 25.5 | 5.7 | 13.6 | 42.9 |
| Computer Network Architects | 5.2 | 20.5 | 32.5 | 6.0 | 14.7 | 42.5 |
| Computer Systems Analysts | 3.7 | 33.5 | 14.2 | 4.4 | 23.7 | 26.7 |
| Network and Computer Systems Administrators | 4.3 | 29.6 | 22.8 | 4.5 | 20.2 | 25.4 |
| Software Quality Assurance Analysts and Testers | 3.7 | 37.9 | 14.9 | 4.1 | 24.6 | 19.7 |

*Notes:* Technical postings with an O*NET code and a parsed minimum experience requirement; the twelve largest occupations by U.S. postings. Junior: 0–2 years; senior: 6 or more.

Table 24: Largest titles in each category, distinct jobs 2019–2025

| Category | Title (normalized) | United States | India | Other three |
|---|---|---|---|---|
| T1 Model building | network engineer | 5,375 | 1,541 | 790 |
| | machine learning engineer | 1,865 | 587 | 241 |
| | senior machine learning engineer | 895 | 314 | 114 |
| | network architect | 617 | 173 | 127 |
| | network operations engineer | 435 | 103 | 172 |
| | network technician | 410 | 18 | 20 |
| T2 Data and ML infrastructure | cloud engineer | 1,851 | 794 | 373 |
| | cloud architect | 996 | 506 | 196 |
| | operations asm | 964 | 2 | 11 |
| | cloud software engineer | 591 | 128 | 91 |
| | lead devops engineer | 386 | 405 | 210 |
| | devops architect | 378 | 331 | 27 |
| T3 Technical application | systems engineer | 9,014 | 895 | 1,235 |
| | senior systems engineer | 4,677 | 608 | 622 |
| | embedded software engineer | 2,865 | 328 | 304 |
| | software engineer iii | 2,122 | 1,338 | 294 |
| | sr systems engineer | 1,750 | 294 | 57 |
| | principal systems engineer | 1,565 | 110 | 177 |
| T5 Functional application | project manager | 27,610 | 6,030 | 9,061 |
| | program manager | 14,493 | 2,156 | 1,353 |
| | business analyst | 11,369 | 7,200 | 4,252 |
| | assistant manager | 9,271 | 4,728 | 999 |
| | machine operator | 8,801 | 36 | 287 |
| | data analyst | 6,585 | 3,393 | 2,356 |
| T4 Product and testing | test engineer | 4,857 | 1,714 | 1,332 |
| | software test engineer | 1,745 | 559 | 300 |
| | qa engineer | 1,509 | 1,893 | 544 |
| | senior test engineer | 1,347 | 1,035 | 425 |
| | java full stack developer | 891 | 1,103 | 55 |
| | user experience designer | 863 | 232 | 157 |
| T6 Governance and security | cyber security engineer | 1,239 | 202 | 168 |
| | cyber security analyst | 1,096 | 577 | 233 |
| | information security analyst | 1,013 | 452 | 204 |
| | security analyst | 926 | 421 | 302 |
| | cybersecurity engineer | 786 | 204 | 61 |
| | cybersecurity analyst | 577 | 91 | 43 |
| T7 General tech | technical support specialist | 2,014 | 374 | 470 |
| | desktop support technician | 1,567 | 61 | 79 |
| | it support specialist | 1,460 | 228 | 320 |
| | it support technician | 1,120 | 63 | 197 |
| | support lead part time | 1,083 | 0 | 53 |
| | technical support representative | 888 | 194 | 293 |

*Notes:* The six largest titles of each category in the fixed lookup, by U.S. distinct jobs. "Other three": United Kingdom, Mexico and the Philippines combined. Titles are shown as normalized (lower case, punctuation removed).

Required experience is parsed with the same English pattern in every country (scripts `extract_experience_required` for the United States and India and `extract_experience_country` for the other countries, in the replication code). A Spanish variant adds under one percentage point of coverage in Mexico. Distinct jobs keep the minimum requirement across their copies.

# H The fixed title lookup

The five-country estimates were first made with an earlier measure, a fixed lookup that gives each job title one category. It is kept as a robustness check on the full-text categories of Section 4, because it labels a job by its title alone and therefore identically in every country and year. This appendix describes it and its known flaws.

The lookup's labels come from an earlier version of the measure, which grouped U.S. and Indian postings by the text of their title and description, year by year, and named each group by its distinctive vocabulary. Because those groups were refit every year, a title could change category between years for reasons that have nothing to do with the job.

I therefore fix the classification at the level of job titles. A title is assigned to a category if it has at least 30 U.S. postings and at least 60 percent of them carry that category's label, pooling 2019–2025. The lookup contains 20,744 focal-category titles and 3,116 lookup *general tech* titles; no title appears in both. It labels a job identically in every year and every country, and it does not depend on how a posting's requirements are written, which matters because Indian and U.S. postings describe requirements differently. The cost is coverage. The lookup holds the title-identifiable core of each category—between 16 and 37 percent of each category's U.S. postings—and matches 3 to 8 percent of each country's postings (United States 4.1, India 7.5, United Kingdom 3.1, Mexico 3.1, Philippines 4.6). Generic titles such as "analyst" or "engineer" are split across categories and drop out.

Appendix Table 24 lists the six largest titles in each category. No category is dominated by one title outside lookup *model building*: the largest title holds between 4 and 9 percent of its category's jobs, and the ten largest between 20 and 30 percent. Most titles are what the category names lead one to expect: machine-learning engineers in lookup *model building*, cloud and DevOps engineers in lookup *data and ML infrastructure*, systems and embedded engineers in lookup *technical application*, test and QA engineers in lookup *product and testing*, security engineers and analysts in lookup *governance and security*, and IT support in lookup *general tech*. Four features need stating plainly.

First, lookup *functional application* is broad. Its largest titles are project manager, program manager and business analyst, and it accounts for 63 to 79 percent of focal-category

jobs in every country. These are roles that put technical systems to work in the business, not jobs defined by AI, and the paper's results for Application should be read with that in mind.

Second, lookup *model building* contains network engineering. Titles containing "network" account for 44.6 percent of its jobs: 45 percent in the United States, 40 percent in India and 78 percent in the Philippines. Network engineer is the category's largest title (21 percent of its jobs), with machine learning engineer second. The earlier labels evidently grouped network infrastructure with model-training work in some years, and the title lookup inherits that grouping.

Third, a small number of non-technical titles enter through firm-specific postings, chiefly assistant manager and machine operator in lookup *functional application*. Titles of this kind account for 2.5 percent of focal-category jobs.

Fourth, and most important for interpretation, a title's category describes the distinctive vocabulary of its postings, not the AI content of the jobs. Read posting by posting, most lookup jobs in every category describe ordinary technical work in which AI is at most one tool among several. The lookup is best read as a classification of technical work by its position in the stack—making models and their infrastructure, building and running software, managing systems—and it is used here only where titles are all that can be compared across five countries.

Table 25 and Figure 21 estimate Equation (5) on these titles: firm × country × title × year cells with the lookup category as $x_z$. Table 5, Panel C, sets their Building–Application differences beside the full-text ones.

# I How postings are assigned to categories

This appendix gives the formulas behind the assignment described in words in Section 4.3.

The assignment has four steps, each a fixed formula; nothing is estimated on the outcome data. Figure 22 summarizes them.

*Windows.* Posting $j$ has a title $\tau_j$ and a description of $N_j$ words. The description is cut into consecutive windows of 350 words,

$$\omega_{jw} = \big(\text{words } 350(w-1)+1,\ \ldots,\ \min(350w, N_j)\big), \qquad W_j = \min\big(\lceil N_j/350 \rceil,\, 8\big), \quad (25)$$

for $w = 1, \ldots, W_j$, so at most the first 2,800 words are read. A posting with no description has one window, its title.

*Embedding.* A sentence-embedding model (BGE-large; Xiao et al. 2023) maps a text to a

Table 25: Focal task categories relative to general tech (T7), by destination: title-lookup categories

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| *Building* | | | | |
| T1 Model building | 0.13 | −0.52* | −0.62** | 0.02 |
| | (0.15) | (0.30) | (0.28) | (0.11) |
| T2 Data and ML infrastructure | 0.35** | −1.04*** | −0.33 | 0.04 |
| | (0.16) | (0.27) | (0.28) | (0.12) |
| *Application* | | | | |
| T3 Technical application | 0.49*** | −1.10*** | −0.40 | −0.05 |
| | (0.16) | (0.29) | (0.28) | (0.12) |
| T4 Product and testing | 0.72*** | −0.51** | 0.30 | 0.06 |
| | (0.14) | (0.26) | (0.29) | (0.11) |
| T6 Governance and security | 0.11 | −0.37 | −0.63** | 0.14 |
| | (0.16) | (0.24) | (0.30) | (0.11) |
| T5 Functional application | 0.28** | −0.16 | −0.04 | 0.22*** |
| | (0.14) | (0.26) | (0.28) | (0.08) |
| *Grouped (separate regression)* | | | | |
| Building | 0.26* | −0.79*** | −0.43 | 0.03 |
| | (0.15) | (0.25) | (0.27) | (0.11) |
| Application | 0.35*** | −0.27 | −0.04 | 0.17** |
| | (0.13) | (0.25) | (0.28) | (0.08) |
| Building − Application | −0.10 | −0.53*** | −0.39*** | −0.14* |
| | (0.10) | (0.15) | (0.11) | (0.08) |

*Notes:* PPML of Equation (5) with destination × category interactions. Reference group: lookup *general tech* (T7; 3,116 titles of networks, servers, systems administration and IT support). 1,864,390 cells; fixed effects firm × title × year and firm × country × year. Standard errors clustered by firm in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The main estimates on full-text categories are in Table 5.

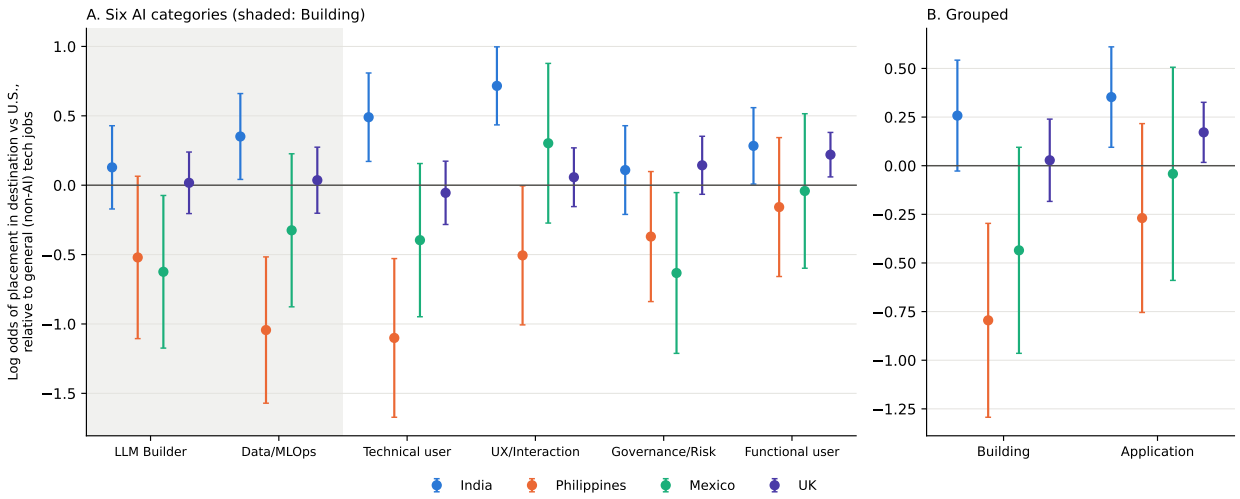



Figure 21: Title-lookup categories: placement of each focal task category in the destination rather than the United States, relative to general tech (T7). Panel A: six categories (shaded: Building). Panel B: Building and Application grouped. 95 percent intervals; standard errors clustered by firm.

unit vector $f(\cdot) \in \mathbb{R}^{1024}$. Each window is embedded together with the title, and the posting's vector is the normalized mean of its windows:

$$e_j = \frac{\bar{e}_j}{\|\bar{e}_j\|}, \qquad \bar{e}_j = \frac{1}{W_j} \sum_{w=1}^{W_j} f\big(\tau_j \oplus \omega_{jw}\big), \tag{26}$$

where $\oplus$ joins the title and the window. A sentence-embedding model is a neural network trained to place texts with similar meaning close together; two postings that describe adapting large language models end up close even if one says "LLM fine-tuning" and the other "customizing foundation models," a match no keyword list written in advance would make. An economist may read $e_j$ as a high-dimensional vector of hedonic characteristics of the job, learned from how language is used rather than chosen by the researcher.

*Centring.* These models place all texts in a narrow cone, so raw similarities are high for every pair. I subtract the corpus mean,

$$\tilde{e}_j = \frac{e_j - \mu}{\|e_j - \mu\|}, \qquad \mu = \frac{1}{|\mathcal{J}|} \sum_{j \in \mathcal{J}} e_j, \tag{27}$$

where $\mathcal{J}$ is the 9,817,530 U.S. and Indian postings of 2019–2025 with a description. The same $\mu$ is used for every country and year, including the five-country postings of Section 4.5. Each category sentence $q_k$ (Section 4.2) is embedded and centred the same way: $\tilde{p}_k = (f(q_k) - \mu)/\|f(q_k) - \mu\|$.

*Assignment.* The posting goes to the category whose sentence is nearest in cosine similarity,

$$k(j) = \arg \max_{k \in \{1,\ldots,7\}} \tilde{e}_j' \, \tilde{p}_k, \tag{28}$$

a nearest-neighbour rule against seven reference jobs. Building and Application are the unions $B = \{1, 2\}$ and $A = \{3, 4, 5, 6\}$; general tech is $k = 7$. The shares in Table 1 are $S_{ckt} = \sum_j \mathbb{1}[c_j = c, t_j = t, k(j) = k] / \sum_j \mathbb{1}[c_j = c, t_j = t]$ over postings. A distinct job (Section 3.3) takes the category of its earliest posting.

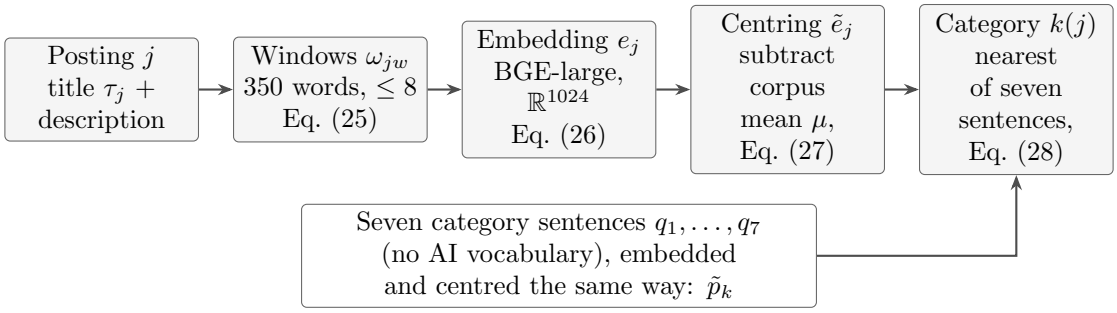



Figure 22: From a posting to a task category. Every step is a fixed formula; nothing is fitted to the outcome data. The same vector $\mu$ and the same seven sentences label postings in every country and year.

*Accuracy.* Let $\mathcal{G}_k$ be the test postings of Section 4.4 whose title puts them in category $k$. Accuracy is the macro average of the per-category hit rates,

$$\text{Acc} = \frac{1}{7} \sum_{k=1}^{7} \frac{1}{|\mathcal{G}_k|} \sum_{j \in \mathcal{G}_k} \mathbb{1}\big[k(j) = k\big], \tag{29}$$

so that a large category cannot carry the average.

*Correcting shares for misclassification.* Write $M_c$ for country $c$'s title-blind confusion matrix, with $M_c[k, l]$ the share of true-$k$ test postings labelled $l$. Observed shares $\hat{S}_c$ relate to true shares $S_c$ by $\hat{S}_c = M_c' S_c$, so the corrected shares are $S_c = (M_c')^{-1} \hat{S}_c$, with negative entries set to zero and the shares rescaled to sum to one; a category with fewer than ten test postings in a country is left uncorrected (its row of $M_c$ is set to the identity). Correcting each country's shares this way widens India's Building shortfall from 6.4 to about 16 percentage points. The correction applies to pooled shares of U.S. and Indian postings, using confusion matrices from clear-title postings; it does not carry over to the within-firm PPML coefficients, whose bias also depends on how errors are spread across firms and years. Appendix K carries the errors into those coefficients by multiple imputation.

*How clearly postings belong.* Figure 23 measures, for the postings of Figure 4, the cosine similarity to the nearest category sentence minus that to the second nearest. The median

margin is smallest for model building (0.027) and data and ML infrastructure (0.039), which sit close to each other, and largest for general tech (0.064); 23.7 percent of postings have a margin below 0.02, and these are where misclassification is most likely. Errors between the two Building categories leave the Building–Application comparison unchanged.

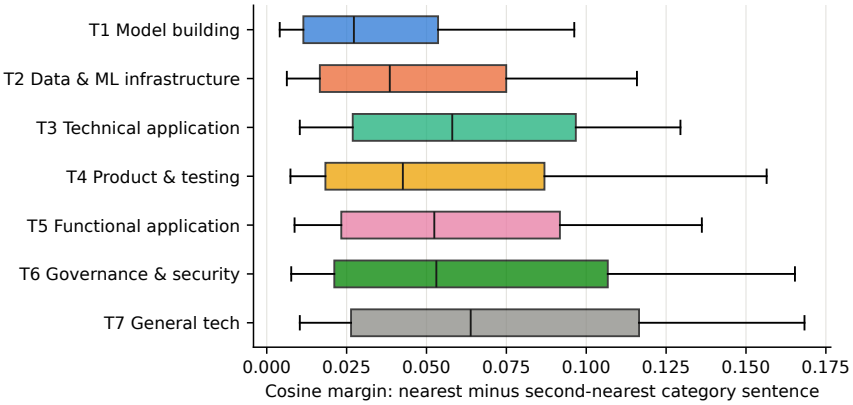



Figure 23: How clearly postings belong to their category. Cosine similarity to the nearest category sentence minus that to the second nearest, by assigned category; boxes span the 25th to 75th percentiles, whiskers the 10th to 90th. Sample of Figure 4.

# J Validating the seven categories against job titles and occupations

A useful category measure should agree with independent signals of the work where those signals are informative, and should add information where they are not. This appendix checks both for the full-text categories of Section 4.2, on the 9.1 million U.S. and Indian technical postings of 2019–2025 that carry one. The full-text encoder reads the title with each window of the description, so agreement with the title is partly mechanical; the known-title checks of Section 4.4 without the title (73.9 percent) and with its phrase also deleted (59.7 percent) are its complement.

Table 27, Panel A, compares the full-text category with the label of the same posting from the fixed title lookup used for the five-country results. Agreement on the seven categories is 38 percent in the United States and 46 percent in India, and on the three groups (Building, Application, general tech) 69 and 80 percent; Cohen's kappa is 0.23 and 0.30, fair agreement. The disagreement has one main source. The lookup places a large share of every category in functional application, the broadest category (it holds 63 to 79 percent of lookup jobs; Section H): 48.5 percent of the U.S. postings that the full text reads as model building carry

Table 26: One title, two jobs: "Data Scientist" at JPMorgan Chase, 2024

| *United States* — "Data Scientist – Vice President," Jersey City, NJ<br>Category: model building (T1) | *India* — "Data Scientist Associate Sr"<br>Category: data and ML infrastructure (T2) |
|---|---|
| "Develop state-of-the art machine learning models to solve real-world problems and apply it to tasks such as NLP, personalization, or recommendation systems." | "Writing SQL Codes/Stored procedures/Python code capturing all the logics to identify the exceptions and testing data, developing Tableau dashboards." |
| "Develop end-to-end ML/AutoML/AutoNLP pipelines and operationalize the end-to-end orchestration of the ML models." | "Ensure codes are designed efficiently to maximize performance and consistency of reporting environment. Enhance ad-hoc reporting capabilities." |
| "Solid background in NLP, Generative AI ... familiar with large language models." "Extensive experience with machine learning and deep learning toolkits (e.g.: Transformers, Hugging Face, TensorFlow, PyTorch ... )." | "Hands-on experience with standard data analytics tools and technologies (SQL, Tableau, Alteryx, Python & Snowflake)." |
| "Published research in areas of Machine Learning, Deep Learning or Reinforcement Learning at a major conference or journal" (preferred). | "Interact with the stakeholders to understand business needs ... Providing regular updates to key stakeholders and senior leaders." |

*Notes:* Verbatim excerpts from the two postings as published on the employer's careers site and collected by LinkUp (U.S. job ID 210494314, posted February 28, 2024; India job ID 210522349, posted June 14, 2024). Both are listed by the employer under the job family "Predictive Science." Selected for clarity, not as representative of a country difference (see Section 4). Categories are the full-text labels of Section 4.3.

a lookup title in functional application, typically a generic analyst or manager title. Apart from that catch-all, each full-text category's most common lookup counterpart is its own category in the United States, and in India for every category except technical application, whose lookup titles split between technical application and product and testing (Panel B). The lookup categories are therefore a coarser version of the same classification, reliable for the Building–Application–general tech split and less so for the finer categories, which is why the U.S.–India results use the full text.

Panel C relates the full-text category to the posting's O*NET occupation. The two are related (Cramér's V of 0.50 in the United States and 0.48 in India), and they agree where they should: 86 percent of U.S. information-security analysts are in governance and security, 79 percent of computer user support specialists in general tech, and data scientists divide between model building (42 percent) and data and ML infrastructure (32 percent). But the occupation does not determine the category. Assigning each posting its occupation's most common category reproduces the full-text category for only 53 percent of U.S. and 59 percent of Indian postings, and in 73 and 83 percent of the occupations with at least 1,000 postings no category holds 80 percent of them. Software developers, the largest occupation, are 48.5 percent technical application in the United States and 65.4 percent in India, with the rest spread over model building, data infrastructure, AI product and testing and general tech. The categories carry information about the work that the occupation code does not.

# K Further checks on the labels, the release test and the skill gradient

This appendix documents the checks cited in Sections 4.4, 6.3, 8.1 and 8.2.

*Misclassification in the coefficients.* Table 29 reports the Building–Application gap when each job's category is redrawn from its country's title-blind confusion matrix. Every variant that does not invert the matrix widens the gap in the three low-wage destinations; only the adversarial swap narrows India's, and only to −0.17. The United Kingdom's gap moves from −0.01 to −0.18 to −0.20 under its own and India's matrices, and stays within 0.05 of zero under the U.S. matrix and the swap.

*Accuracy by country and period.* Table 30 repeats the clear-title check of Section 4.4 for each of the five countries and, for the United States and India, before and after November 2022.

*Label persistence.* Half-years start in November and May, so one boundary falls at the release. The unit is a firm × title × country cell that posts in two consecutive half-years,

Table 27: The full-text categories against the title lookup and O*NET occupations

*A. Full-text category against the title lookup of the same posting*

| Comparison | Country | Postings | Agree (%) 7 categories | 3 groups | Cohen's $\kappa$ |
|---|---|---|---|---|---|
| Fixed title lookup | United States | 1,839,883 | 37.9 | 69.0 | 0.23 |
| | India | 426,471 | 45.6 | 79.5 | 0.30 |

*B. Lookup category of each full-text category's titles (%)*

| | Same category | | Functional application | |
|---|---|---|---|---|
| | U.S. | India | U.S. | India |
| Model building | 17.7 | 25.2 | 48.5 | 45.3 |
| Data and ML infrastructure | 28.1 | 30.4 | 37.7 | 53.2 |
| Technical application | 40.7 | 24.0 | 18.0 | 34.5 |
| Functional application | 87.9 | 96.8 | — | — |
| Product and testing | 9.8 | 22.8 | 73.4 | 69.5 |
| Governance and security | 18.7 | 31.8 | 54.7 | 48.9 |
| General tech | 27.4 | 26.7 | 35.8 | 45.4 |

*C. The full-text category and the O*NET occupation*

| | United States | India |
|---|---|---|
| Cramér's V | 0.50 | 0.48 |
| Occupation's modal category correct (%) | 52.6 | 58.7 |
| Occupations with ≥1,000 postings | 45 | 35 |
| of which no category holds 80% (%) | 73.3 | 82.9 |

*Notes:* U.S. and Indian technical postings of the panel firms with a full-text category, 2019–2025 (KPMG, Deloitte and Infosys excluded). Panel A: postings whose title is in the lookup. Three groups: Building (model building, data and ML infrastructure), Application (the other four focal categories), general tech. Panel B: row percentages of the confusion matrix against the title lookup. Panel C: postings with an O*NET code; the modal-category accuracy assigns each posting the most common full-text category of its occupation.

Table 28: How much the task mix of the same title differs between the United States and each destination

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| Matched firm-title-year cells | 58,940 | 4,588 | 10,666 | 39,345 |
| Share of jobs to relabel (%) | 26.7 | 23.2 | 24.2 | 18.0 |
| Placebo, size-matched (%) | 11.5 | 11.6 | 11.6 | 11.5 |
| Excess over placebo (pp) | 15.2 | 11.6 | 12.7 | 6.4 |

*Notes:* For each firm × title × year with distinct jobs in both the United States and the destination, $D = \frac{1}{2}\sum_k |s_k^{US} - s_k^d|$, the share of jobs that would have to change full-text category to make the two countries' task mixes identical; unweighted mean over cells. Placebo: the same statistic between two random halves of U.S. firm × title × year cells, reweighted so that the size of the smaller side matches the cross-country cells. English postings, 2019–2025; KPMG, Deloitte and Infosys excluded.

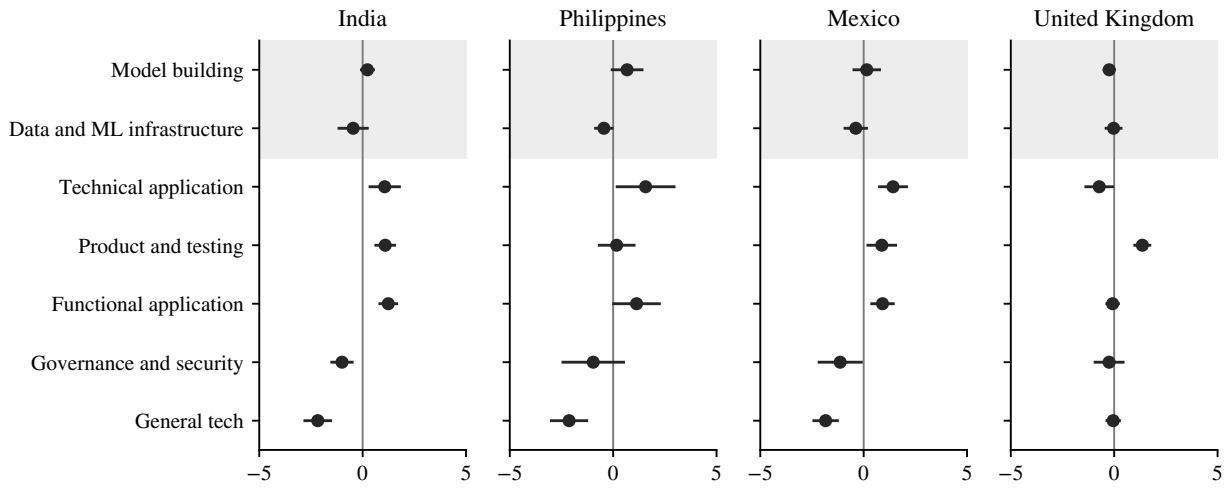



Figure 24: Same title, different work: how a destination's job differs from its U.S. twin. Each point is the coefficient on a destination indicator in a linear probability model of whether a distinct job is in the task category, estimated on the jobs of firm × title × year cells with at least one job in both the United States and the destination, with a firm × title × year fixed effect: the percentage-point difference between a destination job and a U.S. job of the same firm, title and year. 95 percent intervals, standard errors clustered by firm. Shaded: Building. The seven coefficients sum to zero in each panel. English postings, 2019–2025; KPMG, Deloitte and Infosys excluded.

Table 29: The Building–Application gap with misclassification carried into the estimates

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| Observed labels | −0.199 | −0.464 | −0.296 | −0.013 |
| | (0.040) | (0.051) | (0.044) | (0.029) |
| Own matrices | −0.389 | −0.857 | −0.572 | −0.189 |
| | (0.021) | (0.032) | (0.033) | (0.019) |
| Other destinations get India's matrix | −0.388 | −0.810 | −0.533 | −0.175 |
| | (0.021) | (0.033) | (0.030) | (0.018) |
| Other destinations get the U.S. matrix | −0.387 | −0.727 | −0.414 | −0.004 |
| | (0.021) | (0.032) | (0.028) | (0.017) |
| Adversarial: U.S. and India swapped | −0.171 | −0.716 | −0.430 | −0.048 |
| | (0.027) | (0.033) | (0.033) | (0.020) |
| Own matrices, before and after the release | −0.460 | −0.900 | −0.599 | −0.200 |
| | (0.022) | (0.037) | (0.032) | (0.020) |

*Notes:* Building–Application gap of Table 5, Panel B (PPML, firm × category × year and firm × country × year effects, all technical distinct jobs, English-language postings, 2019–2025). Each corrected row averages 40 draws in which every job's category is drawn from $P(k \mid j, c) \propto M_c[k, j]\, s_c[k]$, with $M_c$ the title-blind confusion matrix and $s_c$ the observed shares; standard errors combine within- and between-draw variance (Rubin). U.S. and Indian jobs use their own matrices; the Philippines, Mexico and the United Kingdom share one matrix pooled from their test postings unless stated. A variant that uses the inverted shares as the prior is unstable on India's small test rows (it assigns no true model building) and is not used.

Table 30: Accuracy of the labels on clear-title test postings, by country and period

| | | Title in the input | | | Title and its phrase removed | | |
|---|---|---|---|---|---|---|---|
| | Postings | All | Before | After | All | Before | After |
| United States | 3,883 | 90.2 | 90.2 | 90.8 | 59.6 | 58.7 | 60.2 |
| India | 814 | 91.2 | 91.9 | 91.0 | 58.5 | 59.5 | 61.9 |
| United Kingdom | 137 | 80.7 | 83.5 | 77.9 | 34.3 | 34.5 | 31.3 |
| Mexico | 132 | 83.4 | 63.6 | 80.3 | 57.2 | 40.9 | 52.7 |
| Philippines | 97 | 81.6 | 66.7 | 88.9 | 54.5 | 41.7 | 54.1 |

*Notes:* Macro accuracy (percent; Equation (29)) on test postings whose title leaves the category in no doubt. Before and after: posted before or from November 2022 (postings with no date count in All only). Categories with no test posting in a cell are left out of that cell's average, so the small-country period cells average fewer categories. Title-blind readings were run for the United States and India only.

and persistence is the share of such cells whose modal category is the same in both. The release pair (May–October 2022 against November 2022–April 2023) is compared with the same-season pairs of 2019–2022, with and without role fixed effects. A fall in persistence could also reflect duties that really changed, so it would bound drift from above; there is none.

*Tool fingerprints.* Named tools are extracted from every posting by GLiNER and sorted into three fixed lists: data and model tooling, application languages and frameworks, and infrastructure and support tools; generic AI words are excluded. Each indicator is regressed on the posting's group with firm × title effects, within Software Developers (1.11 million jobs) and within 465 ambiguous titles (0.28 million jobs), with interactions for India and for the period after November 2022.

*The release break against non-technical work.* Table 31, Panel A, replaces general tech by the same firms' non-technical postings, which identifies a shift common to all technical work; Panel B repeats the pooled comparison of the six focal categories with general tech on other units and samples. No version shows a break at the release, and the versions on distinct jobs and on the balanced firms are about twice as precise as the posting version of Table 9.

Table 31: The release break measured against non-technical work, and on other units and samples

| | Break | s.e. | 95% interval | Rank $p$ |
|---|---|---|---|---|
| *A. Against the same firms' non-technical postings, quarterly* | | | | |
| All technical vs non-technical | 0.076 | (0.135) | −0.19 to 0.34 | 0.75 |
| All technical vs white-collar non-technical | 0.004 | (0.111) | −0.21 to 0.22 | |
| General tech vs non-technical | −0.029 | (0.150) | −0.32 to 0.26 | 1.00 |
| *B. Six focal categories against general tech, pooled* | | | | |
| Postings, quarterly (Table 9 design) | 0.095 | (0.123) | −0.15 to 0.34 | |
| Postings, monthly, break November 2022 | 0.109 | (0.118) | −0.12 to 0.34 | 0.82 |
| Distinct jobs, quarterly | 0.010 | (0.057) | −0.10 to 0.12 | |
| Requisitions, quarterly | 0.007 | (0.057) | −0.10 to 0.12 | |
| Firms observed in every quarter | −0.020 | (0.068) | −0.15 to 0.11 | |
| The 510 firms observed in every year | −0.004 | (0.057) | −0.12 to 0.11 | |

*Notes:* Break in India's log odds at 2022Q4 (Panel B, row 2: November 2022) beyond the group's own trend, estimated as in Table 9: PPML on firm × country × group × period counts of U.S. and Indian postings or jobs, 2019–2025, with firm × group × period and firm × country × period effects. Panel A: the groups are technical and non-technical postings (white-collar: SOC major groups 11–29, 41 and 43). Panel B pools the six focal categories against general tech. Rank $p$: share of the true and placebo dates with an absolute break at least as large as the true one (7 quarterly or 21 monthly placebo dates in 2020–2021). Standard errors clustered by firm.

*The PhD gradient.* The nested test compares three Poisson models of the title × category cells of Table 7: M0 with destination slopes on the measured location-cost requirements, M1

adding destination × PhD intensity, and M2 adding unrestricted destination × category effects. M1 improves on M0 (Wald $\chi^2(4) = 94.8$), M2 on M1 ($\chi^2(24) = 411.6$) and PhD intensity on a model with the unrestricted category effects alone ($\chi^2(4) = 105.6$). Research proximity is the share of a title's U.S. postings that match fixed patterns for research teams, applied or fundamental research and publication records. The split-half check assigns each posting to one of two halves by a hash of its identifier, computes the title's PhD share in each half and instruments one with the other in both directions.

# L The O*NET occupations in the sample

LinkUp assigns every posting an O*NET-SOC 2019 occupation code, matched from its title. Table 32 lists every code that appears among the 9,097,904 U.S. and Indian technical postings of the panel firms in 2019–2025 that have a full-text category (KPMG, Deloitte and Infosys excluded): 48 codes, 45 of them in SOC major groups 15 (computer and mathematical) and 13 (business and financial operations), plus Computer and Information Systems Managers (11-3021) and Computer Hardware Engineers (17-2061). Every posting in the frame has a code. For each occupation the table gives its postings in each country and the percentage of its postings, both countries pooled, that Equation (28) places in Building, Application and general tech, with the single most common category.

Three features of the list matter for the paper. The distribution is concentrated: the ten largest codes hold 77.5 percent of U.S. and 83.6 percent of Indian postings, and Software Developers alone 25.6 and 41.9 percent. Most large codes are split across the categories: Software Developers is 14 percent Building, 75 percent Application and 12 percent general tech, and only a handful of codes, such as Database Architects (96 percent Building) and Information Security Analysts (85 percent governance and security), are dominated by one category. And ten codes are mostly Building—Data Scientists, Database Architects, Computer and Information Research Scientists, statisticians and mathematicians among them—but they hold only a third (33.5 percent) of Building postings; two thirds sit inside codes that are mostly something else, Software Developers above all. Section 4 explains why the paper does not use the codes as its task measure, and Appendix J measures how much of the category variation they explain.

Table 32: The O*NET occupations in the sample and their task categories

| O*NET-SOC | Occupation | U.S. | India | Build. | Appl. | Gen. | Modal (%) |
|---|---|---|---|---|---|---|---|
| 15-1252.00 | Software Developers | 1,586,029 | 1,218,032 | 14.1 | 74.5 | 11.5 | Tech. application (56) |

| O*NET-SOC | Occupation | U.S. | India | Build. | Appl. | Gen. | Modal (%) |
|---|---|---|---|---|---|---|---|
| 13-1111.00 | Management Analysts | 926,666 | 351,086 | 9.3 | 84.2 | 6.5 | Funct. application (48) |
| 11-3021.00 | Computer and Information Systems Managers | 528,007 | 169,218 | 19.2 | 55.1 | 25.7 | General tech (26) |
| 15-1212.00 | Information Security Analysts | 426,942 | 102,182 | 2.7 | 89.8 | 7.5 | Gov./security (85) |
| 15-1232.00 | Computer User Support Specialists | 273,176 | 127,417 | 2.6 | 22.4 | 74.9 | General tech (75) |
| 13-1161.00 | Market Research Analysts and Marketing Specialists | 264,249 | 40,899 | 10.4 | 87.9 | 1.8 | Product/testing (57) |
| 15-2051.00 | Data Scientists | 233,218 | 88,177 | 72.7 | 26.5 | 0.8 | Model building (40) |
| 15-1299.09 | Information Technology Project Managers | 222,440 | 96,621 | 13.4 | 70.7 | 15.9 | Product/testing (25) |
| 15-1241.00 | Computer Network Architects | 167,253 | 50,852 | 23.1 | 22.8 | 54.0 | General tech (54) |
| 15-1211.00 | Computer Systems Analysts | 165,215 | 108,228 | 4.0 | 71.1 | 24.9 | Funct. application (48) |
| 15-1244.00 | Network and Computer Systems Administrators | 155,328 | 63,129 | 1.5 | 15.4 | 83.1 | General tech (83) |
| 15-1253.00 | Software Quality Assurance Analysts and Testers | 129,132 | 108,801 | 14.7 | 79.2 | 6.1 | Tech. application (46) |
| 13-1199.00 | Business Operations Specialists, All Other | 126,631 | 43,911 | 6.4 | 77.3 | 16.3 | Funct. application (28) |
| 15-1299.08 | Computer Systems Engineers/Architects | 112,837 | 39,285 | 18.3 | 33.6 | 48.1 | General tech (48) |
| 15-2031.00 | Operations Research Analysts | 106,834 | 37,831 | 12.1 | 68.1 | 19.8 | Funct. application (45) |
| 15-1243.00 | Database Architects | 94,027 | 53,301 | 95.9 | 2.9 | 1.3 | Data/ML infra. (93) |
| 15-1242.00 | Database Administrators | 68,940 | 53,469 | 29.5 | 44.6 | 25.9 | Data/ML infra. (28) |
| 15-1254.00 | Web Developers | 66,340 | 31,196 | 4.1 | 89.7 | 6.1 | Product/testing (70) |
| 15-2051.01 | Business Intelligence Analysts | 65,748 | 24,833 | 15.6 | 82.5 | 1.9 | Funct. application (72) |
| 15-1255.00 | Web and Digital Interface Designers | 63,201 | 12,648 | 0.5 | 99.0 | 0.5 | Product/testing (98) |

| O*NET-SOC | Occupation | U.S. | India | Build. | Appl. | Gen. | Modal (%) |
|---|---|---|---|---|---|---|---|
| 15-1221.00 | Computer and Information Research Scientists | 53,541 | 16,834 | 79.4 | 17.2 | 3.4 | Model building (73) |
| 13-1199.07 | Security Management Specialists | 52,430 | 1,987 | 0.2 | 98.3 | 1.5 | Gov./security (98) |
| 15-1231.00 | Computer Network Support Specialists | 52,199 | 19,805 | 3.8 | 6.0 | 90.2 | General tech (90) |
| 15-2051.02 | Clinical Data Managers | 34,497 | 6,368 | 59.7 | 35.5 | 4.9 | Data/ML infra. (47) |
| 17-2061.00 | Computer Hardware Engineers | 27,278 | 6,132 | 23.7 | 27.6 | 48.7 | General tech (49) |
| 15-2011.00 | Actuaries | 25,289 | 2,935 | 30.2 | 69.0 | 0.8 | Gov./security (46) |
| 15-2041.01 | Biostatisticians | 20,974 | 1,492 | 84.8 | 15.1 | 0.1 | Model building (83) |
| 15-1299.00 | Computer Occupations, All Other | 18,391 | 3,611 | 12.3 | 34.8 | 52.8 | General tech (53) |
| 15-2041.00 | Statisticians | 15,959 | 3,811 | 69.2 | 30.6 | 0.2 | Model building (67) |
| 15-1241.01 | Telecommunications Engineering Specialists | 15,429 | 4,900 | 3.7 | 16.0 | 80.3 | General tech (80) |
| 15-1299.02 | Geographic Information Systems Technologists and Technicians | 14,843 | 1,832 | 38.1 | 46.3 | 15.6 | Model building (25) |
| 15-1299.03 | Document Management Specialists | 12,133 | 2,748 | 4.2 | 87.4 | 8.4 | Gov./security (60) |
| 15-1211.01 | Health Informatics Specialists | 9,890 | 228 | 16.1 | 60.8 | 23.1 | General tech (23) |
| 15-1251.00 | Computer Programmers | 8,993 | 4,894 | 10.8 | 63.5 | 25.7 | Tech. application (37) |
| 15-1243.01 | Data Warehousing Specialists | 7,534 | 4,296 | 86.0 | 11.6 | 2.4 | Data/ML infra. (84) |
| 13-1199.05 | Sustainability Specialists | 7,479 | 773 | 22.5 | 68.0 | 9.5 | Product/testing (24) |
| 13-1161.01 | Search Marketing Strategists | 7,004 | 759 | 22.3 | 76.1 | 1.6 | Product/testing (67) |
| 13-1199.04 | Business Continuity Planners | 5,683 | 891 | 0.9 | 89.9 | 9.2 | Gov./security (78) |
| 15-1299.05 | Information Security Engineers | 3,588 | 753 | 2.2 | 83.9 | 13.9 | Gov./security (80) |

| O*NET-SOC | Occupation | U.S. | India | Build. | Appl. | Gen. | Modal (%) |
|---|---|---|---|---|---|---|---|
| 15-1255.01 | Video Game Designers | 3,036 | 543 | 6.6 | 91.8 | 1.6 | Product/testing (81) |
| 15-1299.01 | Web Administrators | 2,698 | 1,564 | 2.3 | 75.0 | 22.7 | Product/testing (54) |
| 15-1299.04 | Penetration Testers | 1,445 | 249 | 1.9 | 97.8 | 0.3 | Gov./security (91) |
| 15-1299.06 | Digital Forensics Analysts | 1,328 | 42 | 1.8 | 76.9 | 21.2 | Gov./security (73) |
| 15-2099.01 | Bioinformatics Technicians | 1,301 | 46 | 78.9 | 18.0 | 3.0 | Model building (71) |
| 13-1199.06 | Online Merchants | 1,097 | 981 | 0.9 | 97.7 | 1.3 | Product/testing (93) |
| 15-2021.00 | Mathematicians | 712 | 131 | 75.0 | 23.5 | 1.5 | Model building (74) |
| 15-2099.00 | Mathematical Science Occupations, All Other | 664 | 169 | 61.8 | 35.1 | 3.1 | Model building (59) |
| 15-1299.07 | Blockchain Engineers | 312 | 74 | 12.4 | 21.5 | 66.1 | General tech (66) |
| All coded postings | | 6,187,940 | 2,909,964 | | | | |

*Notes to Table 32:* Postings with a full-text category, 2019–2025, KPMG, Deloitte and Infosys excluded. "Build.", "Appl." and "Gen." are the percentages of the code's U.S.-plus-India postings in Building (T1–T2), Application (T3–T6) and general tech (T7); they sum to 100 in each row. Modal category: the most common of the seven categories, with its percentage. Codes are sorted by U.S. postings.

# M Measured location-cost attributes

The main text measures tasks by their category. This appendix reports a complementary measure: four requirements that U.S. employers write into their postings and that tie a task to the United States (security clearance or citizenship, client contact, on-site work, travel). Entered in Equation (9) as the attributes $x_z$, they give the destination-specific prices $\beta_d^\tau = -\theta\delta_d$ of Assumption 2 directly, and so a direct test of the single ranking of tasks in Grossman and Rossi-Hansberg (2008).

## M.1 Measuring the cost of performing a task outside the United States

In the task-trade model, a task moves out of the United States when the wage saving exceeds $\tau(z)$, the cost of performing it away from the firm's U.S. teams. The organizational and trade literatures identify the sources of that cost: the tacit knowledge and communication a

task requires, its need for proximity to customers, its need for physical presence, and legal restrictions on where it may be performed (Antràs et al. 2006; Blinder and Krueger 2013). Each of these leaves a trace in what U.S. employers write into their postings, and I measure four of them directly.

- *Legal and security requirements.* A requirement of a security clearance, U.S. citizenship, ITAR compliance or public-trust eligibility means the task cannot legally be performed outside the United States. This is the purest form of location cost.
- *Client contact.* Client-facing work, direct work with customers, client meetings and stakeholder management require communication that is costly across distance and time zones, and proximity to the clients themselves.
- *Physical presence.* On-site, in-office or in-person requirements tie a task to a facility: hardware, laboratories, secure premises or teams that work face to face.
- *Travel.* A requirement to travel ties the task to a U.S. base from which clients or sites are visited.

For each title in the lookup I compute, from 2.07 million U.S. postings with full text, the share that mention each requirement (the terms are listed in Appendix G). On average 11.4 percent of a title's U.S. postings mention a clearance or citizenship requirement, 13.5 percent on-site work, 11.4 percent client contact and 41.1 percent travel. The composite *location-cost index* is the mean of these four standardized shares, minus the standardized share of postings advertised as remote, standardized across titles. The components are nearly uncorrelated with one another ($|r| \leq 0.16$), so the index aggregates distinct sources of location cost rather than one source measured four times.

Two features of the construction matter for interpretation. Only U.S. text enters, so a destination's own postings—and the way they are written—cannot drive its score. And the index is defined for job titles, so it varies within occupations. Among the titles with the highest scores are cleared government and defense engineering roles and client-facing strategy roles; among the lowest are remote biostatistics, technical-support and product-management roles.

For comparison I also construct a Blinder–Krueger-style occupation index from O*NET Work Context: minus a standardized composite of face-to-face discussion, dealing with the public, physical proximity and public speaking, assigned to each title through its modal O*NET code.

## M.2 Location costs and task location

Table 33 reports the central result. Within the same firm, title and year, a one-standard-deviation increase in location cost lowers the log odds that a job is placed in India rather than the United States by 0.26 (s.e. 0.05), or 23 percent in odds. The Philippines points the same way ($-0.15$, $p = 0.07$). Mexico and the United Kingdom show no gradient, which Section M.3 explains. Estimated for India against the United States alone, with firm × country, firm × decile, country × year and decile × year effects on deciles of the index, the gradient is $-0.54$ (s.e. 0.11). The relation does not reflect occupation mix. India's jobs score 0.032 standard deviations lower than the same firm's U.S. jobs in the same year ($t = -4.5$), and the difference is unchanged when O*NET occupation fixed effects are added.

Table 33: Location cost and task location

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| Destination × location cost (SD) | $-0.258^{***}$ | $-0.151^{*}$ | $-0.035$ | 0.031 |
| | (0.051) | (0.082) | (0.072) | (0.049) |
| Firm × title × year FE | Yes | | | |
| Firm × country × year FE | Yes | | | |
| Cells used | 1,013,952 | | | |

*Notes:* One PPML regression of Equation (5), the four destination interactions estimated jointly. Outcome: distinct jobs in firm × country × title × year cells, 2019–2025. Location cost: composite index from U.S. posting text (Section M.1), standardized across titles. Standard errors clustered by firm in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

The occupation-level index fails the same test, and the reason is instructive. With the Blinder–Krueger-style index, the within-firm gradient for India is $+0.056$ ($p = 0.04$); it falls to $-0.013$ (s.e. 0.058) once Software Developers, about 30 percent of jobs, are excluded. Personal contact varies little across technical occupations, and one occupation drives the index. The text-based measure succeeds because the relevant variation lies within occupations.

## M.3 Destination-specific prices of the attributes

Table 34 enters the separate requirements jointly, and Figure 25 plots them. India avoids every form of U.S.-tied task: client-facing ($t = -3.5$), on-site ($t = -3.4$) and clearance-requiring work ($t = -2.2$). The Philippines avoids client-facing work. Mexico shows no aversion to on-site or client-facing work. The United Kingdom receives *more* client-facing ($t = 2.8$) and on-site work ($t = 3.2$), and less work requiring U.S. clearance or citizenship ($t = -3.3$).

Table 34: Which requirements keep work in the United States, by destination

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| Client-facing | −0.134*** | −0.201*** | −0.091 | 0.087*** |
| | (0.039) | (0.071) | (0.062) | (0.031) |
| On-site | −0.122*** | −0.057 | 0.002 | 0.097*** |
| | (0.035) | (0.051) | (0.033) | (0.030) |
| Clearance / citizenship | −0.099** | −0.007 | 0.063 | −0.224*** |
| | (0.044) | (0.060) | (0.048) | (0.068) |
| Remote | −0.169** | −0.068 | 0.029 | 0.018 |
| | (0.070) | (0.086) | (0.051) | (0.038) |
| Building (0/1) | −0.089 | −0.571*** | −0.434*** | −0.099 |
| | (0.098) | (0.144) | (0.112) | (0.076) |

*Notes:* One PPML regression of Equation (5) with all 20 destination × attribute interactions. Requirements are title-level shares of U.S. postings, standardized; Building indicates lookup *model building* or lookup *data and ML infrastructure* titles. Fixed effects and sample as in Table 33. Standard errors clustered by firm in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The cross-destination contrasts follow the destinations' profiles closely. The United Kingdom receives 0.22 log points more client-facing work than India (s.e. 0.04, $t = 5.5$) and 0.22 more on-site work ($t = 6.3$), with similar gaps against the Philippines and Mexico: client-facing work goes where the clients are. Mexico receives 0.12 log points more on-site work than India ($t = 3.0$), consistent with nearness and shared working hours, though not significantly more client-facing work. Each destination thus receives the tasks whose location costs it is best placed to absorb. A model in which location cost belonged to the task alone could not generate this pattern.

One coefficient runs against intuition: India avoids titles that are frequently advertised as remote in the United States ($t = -2.4$). One reading is that U.S. firms fill remote-capable roles with remote domestic workers, a substitute for placing the work in another country that the pandemic made available (Dingel and Neiman 2020). I note it but do not test it.

# N Reconciling the India estimates

Three ways of counting give three estimates of India's Building–Application difference: −0.42 (posting counts, single years 2019 and 2025: −0.43 and −0.42), −0.15 (full-text labels, five countries, distinct jobs in lookup titles) and −0.09 (title-lookup labels, Table 25). Table 35 moves from the first to the last one change at a time. Every row is a PPML regression with firm × category (or title) × year and firm × country × year effects, 2019–2025, standard errors clustered by firm.

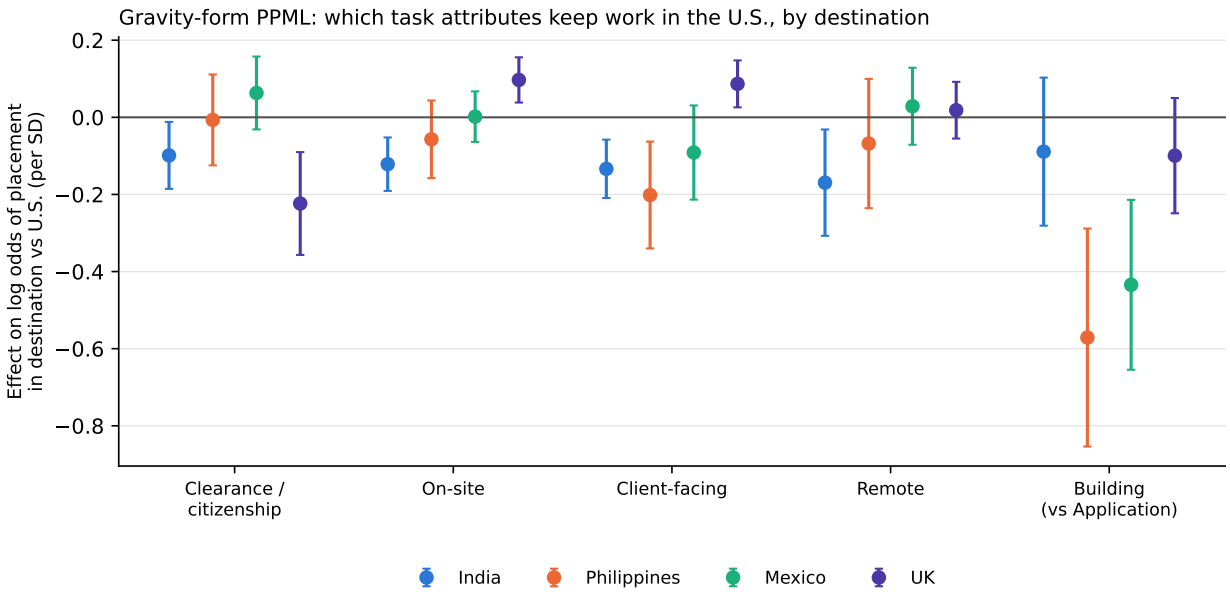



Figure 25: Effect of each requirement on the log odds of placing a job in the destination rather than the United States. Coefficients and 95 percent intervals from Table 34; requirements per standard deviation, Building as a 0/1 indicator.

Table 35: India's Building–Application difference, one change at a time

| Step | Change from the row above | B − A | s.e. |
|---|---|---|---|
| S1 | Full text, U.S.–India, postings, one estimate per year | −0.53 to −0.22 | |
| S2 | Pool 2019–2025 | −0.350 | (0.084) |
| S3 | Keep postings whose title is in the lookup frame | −0.349 | (0.087) |
| S4 | Count distinct jobs (firm × country × title × day) | −0.139 | (0.056) |
| S5 | Labels re-encoded for the five-country frame | −0.151 | (0.058) |
| S6 | Add the U.K., Mexico and the Philippines | −0.153 | (0.057) |
| S7 | Title-lookup labels instead of full text | −0.094 | (0.098) |
| S8 | Firm × title × year cells (Table 25) | −0.094 | (0.098) |
| | S4 with distinct jobs counted by month | −0.060 | (0.041) |
| | S6 on all technical jobs (Table 5) | −0.199 | (0.040) |

*Notes:* India × Building minus India × Application, relative to general tech. S1 reproduces the published year estimates; S6 reproduces the five-country full-text estimate on lookup titles (Table 5, Panel C); the last row widens the sample to every technical job, which moves India's difference by −0.05. The S3 → S4 move (+0.21) matches the copying of postings across locations: U.S. Building jobs appear as 2.84 postings each and U.S. Application jobs as 2.59, against 1.66 and 1.98 in India (log double ratio 0.27).

# O Robustness of the placement estimates

Panel C of Table 5 compares the main estimates with three alternatives. Including the Spanish-language postings (15 percent of Mexico's) changes no gap by more than 0.03. Keep-

ing only jobs in lookup titles, with full-text labels, leaves India (−0.15) and the Philippines (−0.42) close to the main estimates but moves Mexico to −0.11 and the United Kingdom to −0.09. The title lookup of Appendix H, which counts jobs in firm × title × year cells and labels them by title (Table 25), gives the same sign everywhere and gaps of −0.53 for the Philippines, −0.39 for Mexico, −0.14 for the United Kingdom and −0.10 (s.e. 0.10) for India; Appendix N shows that the cell definition, not the labels, accounts for India's smaller value. Mexico's lookup gap, which the full-text labels on lookup titles did not reproduce, is close to the estimate on all technical jobs (−0.30), so it is the restriction to lookup titles, about a sixth of Mexico's technical jobs, that produced the small full-text value. The lookup put the Philippines and the United Kingdom on the same ordering (rank correlation 0.94); on full text they are opposed (−0.66), so that comparison does not survive the change of measure, while India's opposition to the United Kingdom does.

The gap is a log odds ratio. Table 36 measures it instead in percentage points of the probability that a job is placed in the destination rather than the United States, on the same jobs. A linear probability model with firm × year effects, which like PPML weights by jobs, gives a Building job a probability of placement in India 3.1 points lower than an Application job of the same firm in the same year (s.e. 0.6), 1.5 points lower in the Philippines (0.4) and 1.3 points lower in Mexico (0.3), and 0.1 points in the United Kingdom (0.2). Counting every firm-year once, the mean firm-year gap is −0.8, −1.9 and −1.5 points in the three low-wage destinations, all significant, and −0.5 in the United Kingdom, significant at 10 percent only. The two scales rank the destinations differently, and the difference is the reason the paper uses log odds. A percentage-point gap grows with how much of the work a destination receives at all: India receives 42.5 percent of the firms' U.S.-plus-India Application jobs, the Philippines and Mexico under 7 percent of theirs, so the same proportional tilt against Building is many more points in India. The log odds ratio is unchanged when the amount of work a firm places in a country rises or falls overall, which is what a comparison across destinations that receive very different amounts of work needs. The unweighted measures also show that the typical Indian firm-year has little gap (median 0.0 points): the gap is concentrated among the firms that place most of the jobs, as Figure 9 shows.

Every coefficient compares a destination with the United States, and the United States cancels from the difference of two of them: $b_{IN,k} - b_{PH,k}$ is the log odds of placing a category-$k$ job in India rather than the Philippines, relative to general tech. Table 37 reports these differences from the joint covariance matrix of the regression of Table 5. Against the Philippines, India receives model building (+0.52), data and ML infrastructure (+0.69), technical application (+0.82) and product and testing (+0.43) far more readily, and functional application and governance and security about equally; the Philippines' specialty in functional

Table 36: The Building–Application gap in percentage points

| | India | Philippines | Mexico | United Kingdom |
|---|---|---|---|---|
| *A. Log odds (Table 5, Panel B)* | | | | |
| PPML | $-0.20^{***}$ | $-0.46^{***}$ | $-0.30^{***}$ | $-0.01$ |
| | (0.04) | (0.05) | (0.04) | (0.03) |
| The same, in points | $-4.8$ | $-2.3$ | $-1.7$ | $-0.1$ |
| *B. Percentage points of the probability of placement in the destination* | | | | |
| Linear probability, job-weighted | $-3.06^{***}$ | $-1.48^{***}$ | $-1.29^{***}$ | $-0.09$ |
| | (0.64) | (0.42) | (0.29) | (0.18) |
| Firm-year mean, unweighted | $-0.77^{**}$ | $-1.88^{***}$ | $-1.48^{***}$ | $-0.52^{*}$ |
| | (0.31) | (0.54) | (0.38) | (0.29) |
| Firm-year mean, $\geq$10 jobs of each | $-1.25^{***}$ | $-1.59^{***}$ | $-1.50^{***}$ | $-0.34$ |
| | (0.30) | (0.40) | (0.28) | (0.24) |
| Firm-year median, unweighted | $+0.0$ | $-1.0$ | $-1.3$ | $-1.2$ |
| Destination's share of Application jobs (%) | 42.5 | 6.6 | 6.8 | 14.3 |
| Firm-years (unweighted mean) | 10,020 | 2,386 | 3,726 | 8,829 |
| Firm-years ($\geq$10 jobs of each) | 5,252 | 1,516 | 2,319 | 4,825 |

*Notes:* Each column uses the firms' Building and Application distinct jobs in the United States and the destination, 2019–2025, English-language postings, full-text categories, in firm-years with at least one technical job in the destination; KPMG, Deloitte and Infosys excluded. Panel A, row 2: the PPML gap converted to a difference in the probability of placement in the destination at the destination's share of Application jobs. Panel B, row 1: job-level linear probability of placement in the destination on a Building indicator with firm × year fixed effects (estimated on firm × year × group cells weighted by jobs). Rows 2–4: the destination's percent of the firm-year's Building jobs minus its percent of the firm-year's Application jobs, every firm-year with both groups counted once; row 3 keeps firm-years with at least 10 Building and 10 Application jobs. Standard errors clustered by firm in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

application is relative to its own general tech, not an advantage over India. India's edge is larger for Building than for Application, so its gap is 0.27 smaller (s.e. 0.05). Each destination's coefficients come from the firms that hire there, about 3,500 for India and 800 for the Philippines, so the last rows of the table re-estimate the comparison on firm-years with a job in both destinations. The India–Philippines difference is unchanged (+0.26, s.e. 0.06, 694 firms); the India–Mexico difference falls from +0.10 to +0.07 (s.e. 0.04), so part of it reflects which firms hire in each country; the Philippines' gap remains larger than Mexico's (−0.16, s.e. 0.06). In the model, India's and the Philippines' doctoral supply are similarly scarce, so India's smaller gap points to a smaller location cost of Building work there, or to its larger supply of technical graduates below the doctorate, rather than to doctoral supply.

Table 37: Destinations compared with each other: differences of the coefficients of Table 5

| | India − Philippines | India − Mexico | Philippines − Mexico |
|---|---|---|---|
| *A. Category, relative to general tech* | | | |
| T1 Model building | $+0.52^{***}$ | $+0.29^{***}$ | $-0.23^{**}$ |
| | (0.11) | (0.08) | (0.12) |
| T2 Data and ML infrastructure | $+0.69^{***}$ | $+0.39^{***}$ | $-0.30^{**}$ |
| | (0.11) | (0.07) | (0.12) |
| T3 Technical application | $+0.82^{***}$ | $+0.55^{***}$ | $-0.27^{*}$ |
| | (0.16) | (0.09) | (0.14) |
| T4 Product and testing | $+0.43^{***}$ | +0.04 | $-0.39^{**}$ |
| | (0.15) | (0.06) | (0.15) |
| T5 Functional application | +0.03 | $+0.14^{**}$ | +0.12 |
| | (0.10) | (0.07) | (0.12) |
| T6 Governance and security | +0.02 | +0.01 | −0.01 |
| | (0.11) | (0.09) | (0.08) |
| *B. Grouped* | | | |
| Building (T1–T2) | $+0.63^{***}$ | $+0.36^{***}$ | $-0.27^{***}$ |
| | (0.10) | (0.06) | (0.10) |
| Application (T3–T6) | $+0.36^{***}$ | $+0.26^{***}$ | −0.11 |
| | (0.09) | (0.04) | (0.09) |
| Building − Application | $+0.27^{***}$ | $+0.10^{**}$ | $-0.17^{***}$ |
| | (0.05) | (0.04) | (0.06) |
| Building − Application, same firm-years | $+0.26^{***}$ | +0.07 | $-0.16^{***}$ |
| | (0.06) | (0.04) | (0.06) |
| Firms hiring in both destinations | 694 | 1,057 | 458 |

*Notes:* Differences of the destination coefficients of Equation (2) (Table 5), estimated jointly; the United States cancels, so each entry is the log odds of placing a job of the category in the first destination rather than the second, relative to general tech (Panel A) or, in the last row, relative to Application. Same sample and fixed effects as Table 5. Last rows: the grouped regression re-estimated on the United States and the two destinations, keeping only firm-years with a technical job in both destinations. Standard errors from the firm-clustered covariance matrix in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

The firm population is an unbalanced panel (Section 3.2). That alone does not bias these estimates: the firm × country × year fixed effects absorb each firm's presence in each

country and year, including LinkUp's coverage of it. The estimates would be biased if the firms that enter or leave coverage differ in how they divide tasks. Entry is concentrated in India, where the number of firms doubles over the sample. I therefore re-estimate Table 5 on firms observed throughout, with two definitions. The first is the 510 firms with a posting in both the United States and India in every year from 2019 to 2025. The second, stricter one is the 465 firms with a distinct technical job in both countries in every year, which is balance on the jobs the regression uses. Table 38 reports the results. No coefficient moves by more than 1.4 standard errors of the full-sample estimate, and the destination pattern is unchanged. The Building–Application gap widens slightly in the three low-wage destinations, from −0.20 to −0.24 in India, from −0.30 to −0.34 in Mexico and from −0.46 to −0.49 in the Philippines, and stays at −0.01 in the United Kingdom. In India, technical and functional application remain the categories placed there most. The firms that move in and out of the panel therefore place slightly more Building work outside the United States than the firms present throughout, so entry and exit attenuate the Building result rather than produce it. The title-lookup version gives the same answer: on the 510 firms and on the 278 with a lookup-title job in both countries every year, India's gap widens from −0.10 to −0.22 and −0.24, and no coefficient moves by more than 1.5 standard errors. The same restriction widens the U.S.–India levels gap on the posting panel with the six-category labels, from −0.19 (s.e. 0.04) to −0.25 (s.e. 0.05). Coefficients that lose significance in the balanced samples do so because the standard errors are larger with fewer firms; the only sign changes are on coefficients close to zero. The balanced firms are also the largest and longest-observed employers, so part of the difference may reflect how such firms divide work rather than bias from entry. These data cannot separate the two.

# P Further designs for junior hiring

A broader design gives each category (Section 4) the month its core task first became doable by a language model (code generation in November 2022 for technical application and testing work, business analysis in March 2023, tool use in June 2023, open-weight models in July 2023, computer-use agents for IT support in October 2024), and compares categories treated at different dates in a stacked event study, with categories not yet treated and governance and security work, which has no such date, as controls. The category sample is much smaller than the occupation-coded one, and the estimates are imprecise: the pooled junior coefficient is −0.01 (s.e. 0.17) in the U.S. arm and +0.25 (s.e. 0.17) in India, and no category's junior coefficient against the never-exposed category is significant. This design neither confirms nor contradicts the occupation result.

Table 38: Task placement on firms observed throughout 2019–2025

| | All firms | Balanced: any posting | Balanced: in sample |
|---|---|---|---|
| *Building − Application, by destination* | | | |
| India | −0.20*** | −0.24*** | −0.24*** |
| | (0.04) | (0.05) | (0.05) |
| Philippines | −0.46*** | −0.49*** | −0.49*** |
| | (0.05) | (0.06) | (0.06) |
| Mexico | −0.30*** | −0.34*** | −0.34*** |
| | (0.04) | (0.05) | (0.06) |
| United Kingdom | −0.01 | −0.01 | −0.01 |
| | (0.03) | (0.04) | (0.04) |
| *India, by category (relative to general tech)* | | | |
| T1 Model building | −0.07 | −0.06 | −0.06 |
| | (0.05) | (0.07) | (0.07) |
| T2 Data and ML infrastructure | 0.10** | 0.07 | 0.07 |
| | (0.05) | (0.05) | (0.05) |
| T3 Technical application | 0.72*** | 0.71*** | 0.71*** |
| | (0.05) | (0.06) | (0.06) |
| T4 Product and testing | −0.10* | −0.06 | −0.06 |
| | (0.05) | (0.06) | (0.06) |
| T5 Functional application | 0.24*** | 0.30*** | 0.30*** |
| | (0.06) | (0.08) | (0.08) |
| T6 Governance and security | −0.36*** | −0.37*** | −0.37*** |
| | (0.07) | (0.09) | (0.09) |
| Firms | 4,502 | 510 | 465 |
| Distinct jobs | 6,089,734 | 3,423,778 | 3,332,697 |
| Cells | 762,265 | 124,810 | 113,925 |
| Cells used in estimation | 282,703 | 85,144 | 79,514 |

*Notes:* PPML of Equation (2) as in Table 5 (all technical jobs, full-text categories, English-language postings, 2019–2025), on three samples. Column 1 is Table 5. Column 2: firms with at least one posting of any kind in both the United States and India in every year 2019–2025. Column 3: firms with at least one distinct job in this sample in both countries in every year. Their jobs in the United Kingdom, Mexico and the Philippines are kept. Fixed effects firm × category × year and firm × country × year; cells used exclude those absorbed by the fixed effects. KPMG, Deloitte and Infosys excluded. Standard errors clustered by firm in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Treated and control work sit inside the same firms, so AI in one category could change how a firm hires for another. Following the partial-population logic of varying treatment saturation, I ask whether the junior hiring of the never-exposed category after November 2022 moves with how much of the firm's pre-2022 technical work sat in the first-treated categories. In the U.S. arm it does not (+0.10, s.e. 0.24), so there is no sign of spillover onto control work, although the test is imprecise.

The full-text categories allow the same triple difference with the category in place of the occupation group: junior and mid-level postings after November 2022 in each of T1–T6, against general tech. In the U.S. arm, junior postings fell in model building (−0.25, s.e. 0.14) and technical application (−0.15, s.e. 0.09), but no category survives the Holm adjustment across the six (adjusted $p \geq 0.42$), and the placebos at November 2021 are null. In the Indian arm no category shows a junior decline. The category design is less precise than the occupation design and neither confirms nor contradicts it. It was pre-specified.

A last pre-specified design replaces the post-release indicator with a continuous dose: the log training compute of the best publicly usable language model in each month, multiplied by the share of each category's work that AI substitutes for and, separately, the share it complements (both from the Anthropic Economic Index), with the firms' non-software engineers as the comparison group. The design fails on its own construction. Across the sixteen category × country cells the two shares are correlated at 0.9994, so their coefficients are not separately identified and come out as mirror images (U.S. arm: +0.61, s.e. 0.48, and −0.63, s.e. 0.47; Holm-adjusted $p = 0.75$). Their sum, which is identified, is −0.015 in the U.S. arm and +0.021 in India, and between −0.09 and +0.04 in every robustness check. The design therefore shows no relation between the growth of AI capability, weighted by exposure, and the junior share of the more exposed categories, and it cannot distinguish substitution from complementarity.

I also tried to instrument how much work the firm places in India, to ask how much of the U.S. junior change runs through it, and to instrument posting-level AI use directly. Both instruments are weak (Appendix Table 17): a shift-share instrument built from occupation-level changes in India's share of postings, and the abundance of public code in the programming languages a posting names. The firms' changes in India's share are not predicted by their occupation exposure, and posting text records too little AI adoption. The paper therefore does not decompose the U.S. effect into relocation and automation.

# Q Details of the pay comparisons

*The two pay measures.* OEWS group wages weight each occupation's mean wage by its share of each group's U.S. jobs in LinkUp, using the full-text categories; 98 to 100 percent of U.S. jobs in the three groups fall in an occupation that OEWS reports. Because OEWS pays one wage per occupation, a Building and an Application job in the same occupation get the same wage. The two groups' weights overlap by 63 percent (the sum over occupations of the smaller of the two weights): occupation wage growth ranges from 9.6 to 29.2 log points, yet these weights allow a Building–Application difference of only 0.7. That the two groups moved together on this measure is therefore largely built in. For WageScape, each title is assigned to the group that holds at least half of the firms' U.S. jobs with that title, which covers 32,415 titles. General tech titles rose about as fast as other jobs. All figures are nominal unless deflated by CPI-U.

*Title-weighted posted pay.* Assigning each title wholly to one group discards the mix of work within it. Weighting each title's pay growth instead by its U.S. LinkUp jobs in each group, so that a title with 60 Application and 40 Building jobs counts in both, gives the same ordering with a smaller gap. On the 3,155 titles with posted pay in both 2019–20 and 2024–25, pay grew less than in all other U.S. jobs by 11.0 log points in Building (95 percent bootstrap interval 7.5 to 14.5), by 7.1 in Application (5.8 to 8.4) and by 6.8 in general tech (5.0 to 8.5), and Building fell behind Application by 4.0 (0.5 to 7.7). Consumer prices rose 21.2 log points between the same periods, so posted pay on these titles fell in real terms by 9.2 log points in Building, 5.3 in Application and 5.0 in general tech, while posted pay in all other U.S. jobs rose by 1.8. Wages actually paid, on the OEWS measure, were flat in real terms from May 2019 to May 2025 (Building +0.6 percent, Application −0.1, white-collar work −1.0).

*Starting pay and the India gradient.* Across the whole U.S. market, one more log point of 2019 posted pay goes with 17 log points less growth by 2024–25 (s.e. 1.8). Across 1,552 titles observed in both 2019–20 and 2024–25, posted pay grew less in titles whose Indian share rose more, by 17.5 log points per unit of India share (s.e. 3.4), and by 8.8 (s.e. 3.4) once starting pay is held fixed. Measured as the model measures it, against the change in the log U.S. share of the title's jobs in all five countries, posted pay grew 6.0 log points more per log point of U.S. share retained (s.e. 1.5), and 5.1 (s.e. 1.4) with cost weights and starting pay held fixed.

*The size of the productivity effect.* The simplest version of Proposition 5 treats each group as a single task and measures its U.S. share against U.S. plus Indian jobs. That share fell from 77 to 53 percent for Application and from 83 to 57 percent for Building, about

0.37 in logs for each, which with $\theta$ between 2.2 and 4.8 gives 8 to 17 log points of pay. The group's aggregate share equals the cost-weighted average of its tasks' shares only if every task's share falls in the same proportion, and this version leaves out the other destinations. The task-level statistic of Section 7 treats each normalized title as a task, weights it by its 2019 U.S. cost (jobs times the title's posted pay) and measures its U.S. share against all five countries: 0.42 for Application and 0.48 for Building. The other destinations add 0.02 to 0.03 to the fall, the aggregation 0.07 for Building and nothing for Application, and the cost weights 0.01, so the simple version slightly understates the statistic. The task-level version covers only titles present in both years, 29 to 37 percent of base-period U.S. cost; treating titles that vanish with a pseudo-count makes the fall larger still (0.50 to 0.61).

# R Firms whose product is AI

The firms studied so far are, for the most part, users of AI: banks, retailers, consultancies and technology companies whose AI work is one activity among many. The sample also contains a small group of firms whose product *is* AI. I group twenty of them by business model, before looking at their postings: three model developers (OpenAI, Anthropic and DeepMind); six providers of AI compute and infrastructure (Nvidia, Cerebras, SambaNova, Graphcore, Anyscale and Databricks); and eleven AI software and platform firms (DataRobot, H2O.ai, Glean, SoundHound, Moveworks, Builder.ai, Instabase, Uniphore, Eightfold, C3 and Tiger Analytics). Several frontier developers that appear in LinkUp, among them Scale AI, Cohere, Mistral, Perplexity and xAI, are not in the sample, because it is restricted to firms advertising in both the United States and India.

These firms post few jobs, and many of their titles are not in the fixed lookup of Section H. This section therefore uses every posting of the firms, in all titles, and reads each posting's duties directly. A posting counts as *model building* when its duties involve training, designing or substantively changing models; as *AI infrastructure* when they involve data and compute systems built to serve models; and as *AI application* when they involve integrating, deploying or using existing AI systems. The rules are applied sentence by sentence and ignore company boilerplate and lists of desired skills. They have not been validated externally, so the estimates in this section are provisional. Counts are *requisitions*: one advertisement text per firm and country, with the same text posted again after more than 30 days counted as a new requisition. This removes the many U.S. postings that repeat one opening across cities.

These firms advertise AI work of a different kind. Model building accounts for 5.8 percent of their requisitions and AI infrastructure for 20.3 percent, against 0.4 and 0.6 percent at

the other firms. Their footprint outside the United States is small and recent. OpenAI's and Anthropic's first Indian advertisements appear in 2025 and DeepMind's in 2024, while the United Kingdom hosts a large share of DeepMind's and Graphcore's model-building requisitions. Because LinkUp began collecting some of these firms' sites only in 2023, entry dates partly reflect the start of collection rather than of hiring.

Table 39 estimates Equation (5) with the three task types as the task dimension, separately for these firms. The firm × task × year and firm × country × year effects play the same roles as in Section 6, and the coefficient of interest is the difference between a destination's coefficient on model building and on AI application. Unlike the average firm in the sample, AI-native firms advertise model building outside the United States less than AI application, in India as well as in the United Kingdom. Within firm and year, the destination-to-U.S. ratio for model building is 31 percent lower than for application in India (−0.38 log points) and 46 percent lower in the United Kingdom (−0.61). A second definition, the 27 firms with at least 500 postings and at least 40 percent AI work among their 2019–2021 postings, gives similar estimates (−0.34 and −0.43). For all other firms the same contrast is close to zero in India (0.06, s.e. 0.14) and −0.26 (s.e. 0.07) in the United Kingdom, in line with the small Building–Application differences of Table 5. The twenty firms are too few to move these estimates.

The two results differ in strength. With twenty firms, conventional clustered standard errors are too small, so Table 39 also reports the range of estimates when each firm is dropped in turn and a bootstrap over firms. The United Kingdom estimate survives the exclusion of any single firm (range −0.69 to −0.47), and the bootstrap interval excludes zero (−0.95 to −0.11). The India estimate is negative whichever firm is excluded, but it is −0.22 without Nvidia, the largest advertiser, and its bootstrap interval (−0.57 to 0.12) includes zero. The India evidence is therefore suggestive. It is consistent with firms that develop models keeping that work close to their U.S. research organisations while placing application and integration work outside the United States, a pattern that does not appear among the much larger group of firms that use AI. The Mexico estimate rests on one or two firms and separates when C3 is dropped, so it is not informative. The Philippines is omitted because the twenty firms post no model-building requisitions there.

Figure 26 shows India's share of U.S.-plus-India requisitions by year. India's share of the AI-native firms' model-building requisitions rose to 28 percent in 2023 and then fell to 8 percent by 2025, as their U.S. model-building requisitions roughly tripled (from 139 to 408). Among the other firms, India's share kept rising, to 35 percent. The decline comes from growth in U.S. model building, not from fewer Indian requisitions (36 in 2025, against 54 in 2023). OpenAI's and Anthropic's first Indian advertisements (8 and 4 requisitions over the

Table 39: AI-native firms: model building versus AI application, destination relative to the U.S.

| | India | United Kingdom | Mexico |
|---|---|---|---|
| Named AI-native firms (20) | $-0.38^{***}$ | $-0.61^{***}$ | $-0.99^{***}$ |
| | (0.08) | (0.17) | (0.21) |
| Pre-period AI-intensive firms (27) | $-0.34^{***}$ | $-0.43^{***}$ | $-0.49^{**}$ |
| | (0.13) | (0.13) | (0.22) |
| All other firms | 0.06 | $-0.26^{***}$ | $-0.27^{*}$ |
| | (0.14) | (0.07) | (0.14) |
| All firms excluding the named AI-native firms | 0.06 | $-0.26^{***}$ | $-0.27^{*}$ |
| | (0.14) | (0.07) | (0.14) |
| Named firms: leave one firm out, range | [-0.41, -0.22] | [-0.69, -0.47] | [-13.40, -0.91] |
| Named firms: firm bootstrap 95% interval | [-0.57, 0.12] | [-0.95, -0.11] | [-14.61, -0.79] |

*Notes:* $\beta_{dA} - \beta_{dC}$ (model building minus AI application) from PPML of requisition counts in firm × country × task × year cells with firm × task × year and firm × country × year fixed effects, estimated separately for each firm group; firm-clustered standard errors in parentheses; $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$. Named AI-native firms: 20 companies whose product is AI models, AI compute or AI software (OpenAI, Anthropic, DeepMind, Nvidia, Cerebras, SambaNova, Graphcore, Anyscale, Databricks, DataRobot, H2O.ai, Glean, SoundHound, Moveworks, Builder.ai, Instabase, Uniphore, Eightfold, C3, Tiger Analytics). Pre-period AI-intensive: 27 firms with at least 500 postings and at least 40% AI work among 2019–2021 postings. With 20 firms, clustered standard errors are unreliable; the bootstrap resamples firms (499 draws). The Mexico estimate rests on very few firms and separates when C3 is dropped. The Philippines is omitted: the named firms post no model-building requisitions there. Task labels are provisional (unvalidated description classifier).

whole period) include no model building. The annual counts for AI-native firms are small.

These comparisons describe where firms advertise, not where they employ people, and they do not identify why AI-native firms differ. Their research staff, compute, data and leadership are concentrated in the United States, and any of these could produce the pattern. So could the fact that their Indian operations are new.

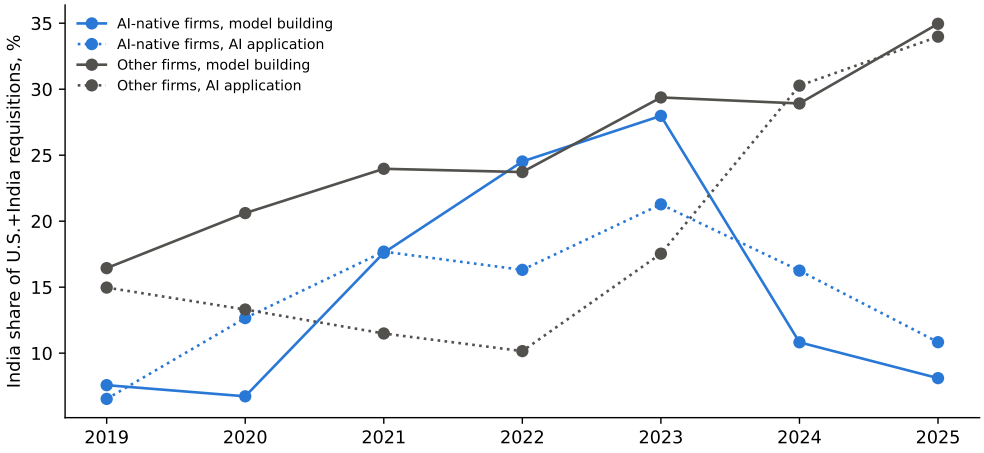



Figure 26: India's share of U.S.-plus-India AI requisitions, AI-native and other firms, by year, for model building (solid) and AI application (dotted). Requisitions labelled from posting descriptions (provisional). Counts for AI-native firms are small: 5 to 64 Indian model-building requisitions per year.