WASHINGTON CENTER FOR
Equitable Growth
Evidence for a Stronger Economy

WORKING PAPER SERIES

What Work Does Generative AI Do?

Alexander Bick
Adam Blandin
David Deming
Tyler Schumacher

View Report Online

September 2026

www.equitablegrowth.org

601 13th Street NW, 12th Floor
Washington, D.C. 20005

(202) 545-6002

What Work Does Generative AI Do?*

Alexander Bick
Federal Reserve Bank of St. Louis & CEPR

Adam Blandin
Vanderbilt University

David J. Deming
Harvard University & NBER

Tyler Schumacher
Vanderbilt University

First version: April 27, 2026
Current version: August 19, 2026

Abstract

We measure how workers use genAI for their jobs in a nationally representative survey linking genAI adoption to detailed occupations and tasks. Our data provide the first task-level genAI adoption indexes, which we show can inform analyses of genAI's labor market impact. Exposure scores explain some, but far from all, of the variation in adoption across occupations and tasks. We also distinguish our indexes from measures based on genAI platform chat logs, which differ conceptually and tend to over-classify chats into generic activities spanning many occupations. Finally, we highlight that current adoption is widespread but shallow: genAI is used across many occupations and tasks, yet within most of them, fewer than half of workers adopt. This indicates substantial variation among workers doing very similar work, suggesting that understanding who adopts may matter as much as understanding which tasks genAI assists.

JEL Codes: J24, O33, C83
Keywords: Generative AI, Technology Adoption, Occupation and Task-Level Measurement, AI Exposure, Labor Markets

* *Contact*: alexander.bick@stls.frb.org; adam.blandin@vanderbilt.edu; david_deming@harvard.edu; tyler.schumacher@vanderbilt.edu. Kevin L. Bloodworth and Khanh Le provided excellent research assistance. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System. This work was supported by the Walmart Foundation, the Harvard Business School AI Institute, and a joint grant from the Washington Center for Equitable Growth and the Russell Sage Foundation.

1 Introduction

Generative artificial intelligence (genAI) tools are rapidly diffusing into workplaces (Bick et al., 2026b). A central question is how this new technology will affect worker productivity, wages, and employment. Prior research on technological change emphasizes that these labor market effects hinge on the type of work a new technology performs (Autor et al., 2003). Tasks performed by a new technology tend to become less valued in the labor market, while complementary tasks that the technology does not do tend to become more valued. Understanding genAI's labor market impact therefore requires first understanding what work genAI does.

We address this measurement need using novel data from the Real-Time Population Survey (RPS), a repeated national online labor market survey that has run since 2020 (Bick and Blandin, 2023). The RPS introduced a genAI adoption module in August 2024, and since August 2025 has measured genAI adoption at the detailed task level.

A key challenge in trying to learn what work genAI does is that it is infeasible to ask a survey respondent about thousands of potential work tasks that they might do. The RPS addresses this challenge in two steps. First, for each worker we elicit their detailed occupation based on the 2018 Standard Occupational Classification (SOC) system. Second, in four survey waves between August 2025 and May 2026, we use the worker's occupation to identify their ten most important work tasks (Detailed Work Activities) according to the Bureau of Labor Statistics' ONET database. We present these ten tasks to respondents and ask them to select the tasks they regularly perform in their job. We then ask which of those tasks they used genAI to help with. This information allows us to construct task-level adoption rates: among workers who do a given task, what share use genAI for it?

Our task-based measures pass several basic diagnostic checks. The share of workers who perform a given task varies widely across tasks and remains stable across survey waves, suggesting that these measures contain meaningful information. Similarly, task-level genAI adoption rates also vary widely across tasks, and their rankings remain stable across survey waves. Finally, task shares in the RPS are similar to an economy-wide benchmark based on occupation shares and occupation-task mappings from ONET; this indicates that restricting attention to workers' ten most important tasks does not distort their relative importance in the US economy.

We use this adoption data to make four contributions toward understanding what work genAI does. First, we construct detailed occupation- and task-level adoption indexes and pro-

vide a simple economic framework showing how these measures can inform the labor market effects of genAI.[1] The indexes reveal substantial differences in adoption across occupations and tasks. Occupations with the highest adoption rates are management and professional occupations, especially business, finance, and computer-related occupations. Tasks with the highest genAI adoption rates involve cognitive and information-intensive work, including data analysis, interpretation, and writing. In contrast, adoption is lowest in personal service occupations and in tasks requiring manual activity or direct interaction with people or equipment. Some of this work also involves care or safety, where responsibility or liability may limit genAI use.

Second, we compare adoption rates with several "genAI exposure scores", which predict which occupations and tasks genAI is most likely to directly impact (Brynjolfsson et al., 2018; Eloundou et al., 2024; Felten et al., 2021; Webb, 2020). Exposure scores predict between 5% – 53% of adoption variation across occupations and 7% – 44% of variation across tasks. The results help establish that these widely used measures capture some, but far from all, of variation in genAI use. In particular, clerical and administrative occupations and tasks tend to have high predicted exposure relative to adoption. These workers may have modest technology skills and face compliance or data privacy barriers that limit adoption. Occupations and tasks with low predicted exposure but high adoption tend to feature high levels of autonomy, which may minimize adoption barriers and provide greater incentives for workers to produce more.

Third, we compare task-level genAI use in the RPS with measures based on chat logs from Handa et al. (2025) (Anthropic Claude), Chatterji et al. (2025) (OpenAI ChatGPT), and Tomlinson et al. (2025) (Microsoft Copilot).[2] The four measures differ substantially, with all pairwise correlations across datasets below 0.4. Much of this disagreement appears to stem from how chats are classified into tasks. ONET tasks combine a basic activity with the occupational purpose it serves, so assigning a chat correctly requires knowing the user's occupation. Chat classifiers observe the text of a conversation but typically not the user's occupation. As a result, they tend to assign chats to a few tasks with generic, activity-based titles. For example, more than 15% of OpenAI chats are classified as "Edit written materials or documents," even though only 2.4% of workers are in occupations that include this task in ONET. In most occupations, editing is instead embedded in higher-order tasks such as preparing reports, drafting legal documents, or writing up research. The tasks most

[1] Adoption indexes can be downloaded from https://sites.google.com/view/covid-rps/generative-ai.

[2] We omit Iscenko et al. (2026) (Google Gemini) from our analysis because their task-level data are not publicly available.

over-represented in chat data follow this pattern, including "Design computer or information systems or applications" and "Gather information from physical or electronic sources." These generic activity tasks act as catch-all categories, resulting in very concentrated chat shares: the top ten tasks account for 46%–61% of genAI use in chat data, compared with 22% in the RPS.

Over-classification of generic activity tasks in chat data, relative to ONET, has two practical implications. First, chat-task classifications have been used to construct genAI occupation shares, which in turn have been used as proxies for genAI exposure across occupations, e.g. Brynjolfsson et al. (2026), Edlich and Slok (2026), Gimbel et al. (2025), and Massenkoff and McCrory (2026). Our results suggest that these proxies are likely inaccurate and offer a direct adoption-based alternative. Second, rather than using detailed ONET tasks, chat data may be better suited to a task framework organized around basic activities. For example, earlier waves of the RPS also identify writing as the most common basic activity assisted by genAI, consistent with the OpenAI chat data (Bick et al., 2026b; Chatterji et al., 2025).

Finally, we highlight that current genAI adoption is widespread but shallow: genAI is used in many occupations and tasks, but few of these exhibit very high adoption rates. For example, four out of five detailed occupations have adoption rates above 20%, but only one out of six occupations exceed 70% adoption. Similarly, two out of five detailed tasks have adoption rates above 20%, while no task exceeds 70% adoption. This implies substantial adoption variation among workers performing very similar work. Exposure scores and chat-platform task shares are largely silent about such worker-level variation within tasks: the former are defined at the task level, and the latter contain no information about non-adopters. But our findings suggest that understanding why some workers adopt and others do not may matter as much for genAI's labor market impact as understanding which tasks genAI does.

We provide evidence consistent with one potential driver of worker-level adoption variation: experience with genAI in one domain appears to promote adoption in others. We show that some workers systematically use genAI for more tasks than others; experienced users adopt genAI for more tasks; workers who perform high-adoption tasks are more likely to adopt genAI for their other tasks; and plausibly exogenous shifters in workplace adoption predict adoption outside of work. These patterns are consistent with a learning mechanism: workers may incur an initial fixed cost to learn how to use genAI for cases with clear benefits, then apply genAI to other domains once that cost is paid. If so, adoption within occupations and tasks will continue to deepen over time as experience with genAI accumulates.

Another paper measuring genAI adoption across occupations is Lindenlaub et al. (2026), who use a representative 2024 survey of German workers linked to administrative records. Like us, they find that exposure scores are only moderately predictive of realized adoption across occupations. They rationalize these gaps with a quantitative structural model, similar in spirit to our stylized framework, in which adoption depends on comparative advantage: occupations with high exposure may still have low adoption if workers remain more productive than genAI or if genAI is costly to use. In contrast to them, we measure adoption at the detailed task level. These measures provide a rich picture of what genAI does and shows that worker-level variation is sizable even among workers performing very similar activities. Our repeated survey collection also allows us assess the stability of these patterns and track changes in adoption over time.

By measuring realized genAI use at the task level, our approach provides a simple and portable input for research on the economic effects of genAI. The task- and occupation-level adoption indexes we construct can be directly incorporated into existing analyses of productivity, wages, task reallocation, and inequality, and can be updated as adoption patterns change. The resulting indexes will help researchers move beyond predictions over technological feasibility toward an understanding of what work genAI actually does.

The remainder of the paper proceeds as follows. Section 2 introduces our primary data source and describes how we measure genAI adoption at the task and occupation level. Section 3 presents our conceptual framework and documents genAI adoption rates across occupations and tasks. Section 4 compares adoption rates to existing genAI exposure predictions. Section 5 investigates individual-level determinants of adoption. Section 6 discusses channels through which genAI may drive future productivity gains. Section 7 compares our survey-based measures to task-level measures constructed from genAI chat logs. Section 8 concludes.

2 Data and Measurement

2.1 The Real-Time Population Survey (RPS)

Our main data source is the Real-Time Population Survey (RPS), a national labor market survey of U.S. adults aged 18-64 (for a detailed discussion, see Bick and Blandin 2023). The RPS is fielded online by Qualtrics, a large commercial survey provider, and has collected multiple survey waves each year starting in 2020. Since August 2024 the RPS has been fielded quarterly.

The RPS is designed to mirror the Current Population Survey (CPS) along key dimensions. RPS responses are collected within two weeks, with either the first or second week overlapping with the CPS reference week. The RPS survey borrows questions on demographics and labor market outcomes from the basic CPS and CPS Outgoing Rotation Group, using the same word-for-word phrasing when practical and replicating the intricate sequence of questions necessary to elicit labor market outcomes in a manner consistent with the CPS (US Census Bureau, 2015). Replicating key portions of an existing high-quality survey ensures that survey concepts are comparable, which allows researchers to validate RPS outcomes against a widely used benchmark with a larger sample size and to construct sample weights.

In August 2024, the RPS added a module measuring individuals' use of genAI. In August 2025, we introduced task-level measures of genAI adoption. The analysis in this paper is based on four survey waves that include these measures, conducted in 08/25, 11/25, 02/26, and 05/26.

2.1.1 RPS Sample

Qualtrics panel members are recruited to the panel online and, in our case, can participate in exchange for 30 to 50 percent of the $5.50 paid per completed survey. The Qualtrics panel includes about 15 million members and the full panel is not a random sample of the U.S. population. However, Qualtrics can target survey invitations to specific demographic groups. The RPS sample was designed to be nationally representative across key demographics, including sex, age, race and ethnicity, education, marital status, household income, and geographic region. We also include measures to filter out low quality responses. Details on sampling and data quality procedures can be found in Appendix A.1.

Each survey round collected responses for about 5,000 individuals. Appendix Table A.1 compares the unweighted sample composition between the CPS and RPS along the demographics targeted in the sampling procedure for our main surveys (columns 1 and 2). The bottom panel also compares employment status in the CPS and RPS, which was not targeted in the sampling procedure. The two surveys yield similar population shares.

2.1.2 Sample Weights and Validation

To address remaining discrepancies, we construct sample weights using the raking algorithm of Deming and Stephan (1940), ensuring that the weighted sample proportions align with the

demographic characteristics targeted in the sampling procedure and that the employment shares in the RPS align with the CPS as well. Details on weighting can be found in Appendix A.2.

Despite being representative based on observable characteristics, our survey samples could still potentially suffer from selection on unobservable characteristics correlated with our variables of interest. However, Bick and Blandin (2023) and Bick et al. (2025a,b) show that the RPS closely aligns with US government household surveys on employment, hours worked, earnings, industry composition, employee tenure, and work from home, none of which were targeted by our sampling scheme. Bick et al. (2026b) show that ChatGPT adoption in the RPS closely aligns with estimates from other surveys (Fletcher and Nielsen, 2024; McClain, 2024). In particular, the latter estimates are derived from a survey using traditional random address-based sampling and includes respondents without internet access. Appendix A.3 shows that the distribution of usual weekly earnings, industry, and college major shares in the RPS, not targeted during sampling, is similar to the CPS.

2.2 Measuring Occupation

Following Miano (2025), we elicited respondents' job titles through a free text response:

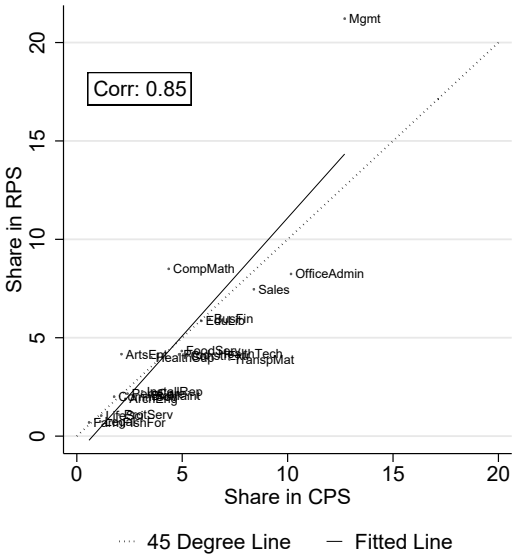
> *What is your main occupation? Please type your occupation in the box below and select one of the suggested options. Try to be specific. For instance, write preschool teacher or high school teacher rather than just teacher. If none of the options corresponds to your occupation, try adding more detail or rephrasing.*

Once respondents have starting typing, they are presented with autocomplete suggestions covering over 40,000 occupations from ONET and the Occupational Outlook Handbook (OOH), following Miano (2025). We extend this method by not forcing respondents to select one of the autocomplete suggestions. Instead, we present these individuals with the five most probabilistic SOC occupations based on their industry and stated job title using a matching algorithm developed by the National Institute for Occupational Safety and Health (Laughlin et al., 2024). We do the same for occupations from the OOH that do not map into a unique SOC. In total, about 40% of workers select an initial autocomplete suggestion. The remaining 60% are routed to the autocoder; among these, about 7% (or 4% of all employed) select "none of the above". We drop these respondents from our sample.

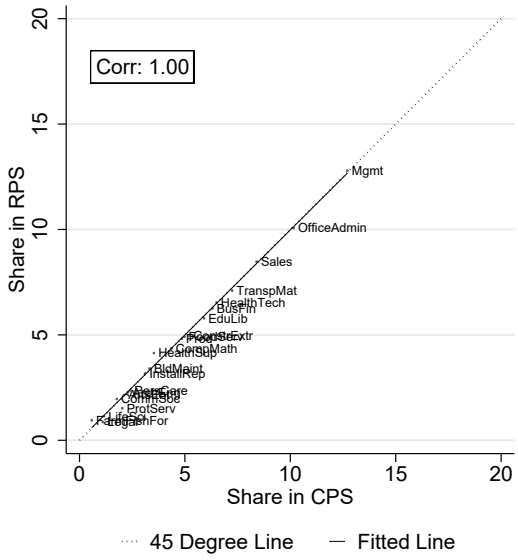
Figure 1 compares the distribution of occupation employment shares in the RPS and

Figure 1: Comparing Occupation Shares in the RPS and CPS

(a) Occupation Shares: Unweighted



(b) Occupation Shares: Weighted



Notes: The figure on the left uses unweighted RPS data and the figure on the right uses weighted RPS data. We use the same sample of RPS respondents in both figures. Data samples for the occupation comparison are employed respondents ages 18-64 in the pooled August 2025, November 2025, February 2026, and May 2026 RPS and CPS with sample sizes of 13920 and 158472, respectively.

CPS. (Occupation shares were not targeted by our sampling scheme.) Figure 1a shows that occupation shares in the unweighted RPS data are overall very similar to those in the CPS. Our weighting scheme includes occupation employment shares from the CPS, and Figure 1b verifies that occupation shares in the weighted RPS data nearly perfectly match shares in the CPS.

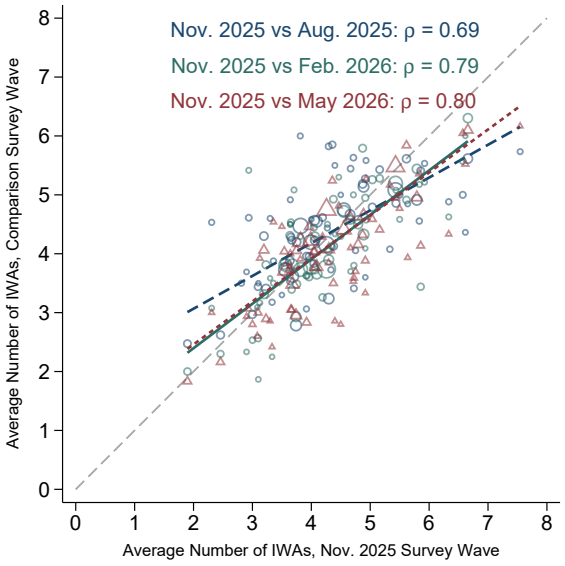
2.3 Measuring ONET Tasks

To learn how workers are using genAI, we ask respondents whether they perform a set of detailed tasks in their job, and whether they use genAI to help them accomplish those tasks. A key practical challenge we face is that it is infeasible to survey respondents about thousands of detailed tasks that they could potentially perform. Our solution is to use a respondent's occupation to focus on the most important tasks associated with their job.

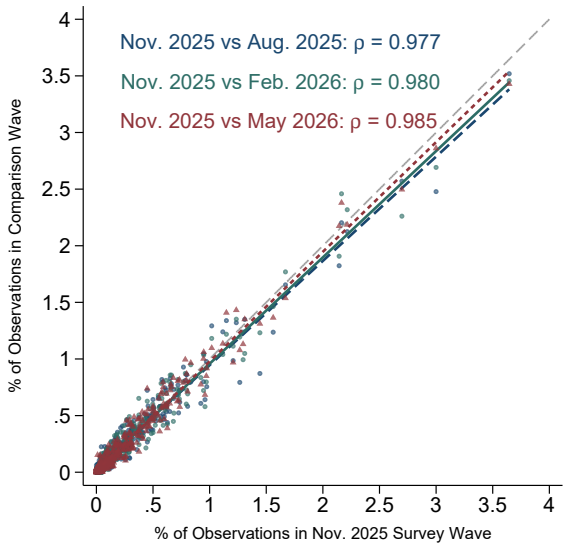
ONET contains two sets of granular activities: Tasks and Detailed Work Activities (DWAs). ONET assigns Importance Scores to each of an occupation's Tasks, which re-

Figure 2: Task Prevalence in the RPS by Survey Wave

(a) Average Tasks by Occupation



(b) Task Shares



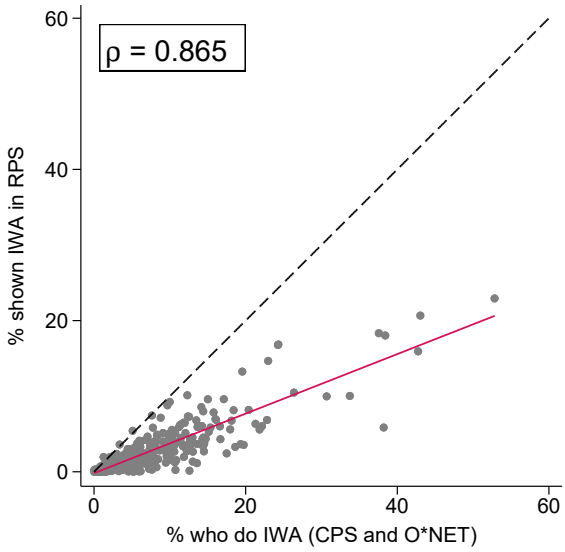
Notes: Panel (a) plots the mean number of IWAs associated with each occupation in the November 2025 survey wave against the corresponding mean in the August 2025 wave, the February 2026 wave, and the May 2026 wave. Each observation is an ONET SOC title with at least 10 observations. Panel (b) plots the share of each task among all tasks performed. Each observation is an IWA. In each figure the dashed gray line is a 45-degree line. The values of ρ report the correlation coefficient between two surveys.

flect how essential a Task is to overall job performance in that occupation. However, we opt to use DWAs because they tend to be more concise and intuitive, making them easier to include in a survey. While DWAs are not directly assigned an Importance Score, ONET provides a mapping from Tasks to DWAs. For DWAs associated with a single Task, we assign the Task's Importance Score to the DWA. When multiple Tasks map into a DWA, we assign that DWA the average Importance Score of its associated Tasks. Details on this procedure are provided in Appendix A.6.

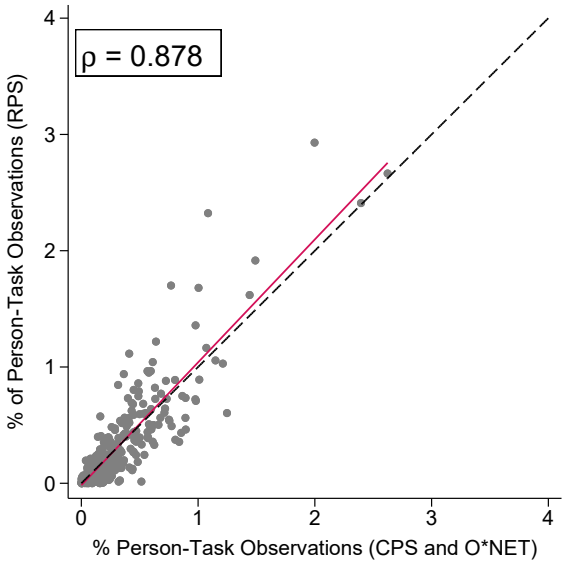
After identifying a respondent's occupation, we present them with the 10 most important DWAs for their occupation. We ask them to select all of the DWAs that they perform as a part of their job; we then ask for which of those DWAs they regularly use genAI. Some of our analysis is conducted at the DWA level. However, the 2,040 DWAs can also be aggregated into 332 Intermediate Work Activities (IWAs), 37 Work Activities (WAs), and 9 Broad Work Activities (BWAs), which we also use. For the remainder of the paper we use the parlance of the literature and refer to ONET work activities as "tasks."

Figure 3: Comparing Task Prevalence in the RPS and CPS

(a) Shares of people shown task



(b) Shares of person-task observations



Notes: The figure on the left shows the share of people shown each IWA in the RPS against the share of all people who, based on CPS employment shares and the associated ONET task mappings, do that task. The figure on the right shows the share of person-task observations the IWA accounts for based on CPS employment shares and ONET task mappings against the share of person-task observations in the truncated RPS.

Figure 2a shows that the mean number of tasks selected varies by occupation. In most occupations, the average number of tasks selected is between three and six. (Across all occupations the average number of tasks selected is 4.7, and 99% of workers report performing at least one task; see Appendix Figure D.1.) Figure 2b shows that tasks also vary in their overall prevalence. Both of these measures are stable across the four separate RPS waves. This indicates that the task estimates are internally consistent and helps to address concerns over excessive measurement error.

A limitation of our task data is that we restrict attention to the top ten tasks for each occupation. To assess the share of work that this approach omits, and whether it presents a biased view of task composition, Figure 3 compares task prevalence (at the IWA level) in the RPS to benchmarks constructed from the CPS and ONET. In Figure 3a, the x-axis shows the share of CPS workers whose occupation contains task t in ONET, while the y-axis shows the share of RPS workers who were shown task t (because it was among the top ten tasks in their occupation). The RPS omits tasks outside the top ten, so by construction

most observations lie below the 45-degree line.[3] However, the correlation between the two measures is 0.865, indicating that task share rankings are broadly similar.

Figure 3b compares economy-wide task shares across the two surveys. For each task t, we calculate the share of all worker-task pairs associated with that task. In the CPS benchmark, we assign each worker all tasks associated with their occupation in ONET and calculate the resulting task shares using the CPS occupation distribution. In the RPS, we repeat the same calculation using only the subset of tasks tasks shown to respondents. The x-axis reports the CPS-based measure and the y-axis the corresponding RPS measure. Intuitively, this comparison asks whether tasks that are more common in the economy according to the CPS and ONET are also more common in the RPS. The two measures are highly correlated ($\rho = 0.878$), and the best-fit line more closely tracks the 45-degree line. This indicates that (i) tasks which are more common according to the CPS and ONET are also more common in the RPS, and (ii) quantitatively, the task shares in the RPS are a fairly close approximation to those in ONET.

From Figures 2 and 3, we conclude that task measurement in the RPS omits a meaningful share of work but still provides a useful approximation to the economy-wide task distribution.

2.4 Measuring Generative AI Use

The RPS genAI module begins with a definition of genAI:

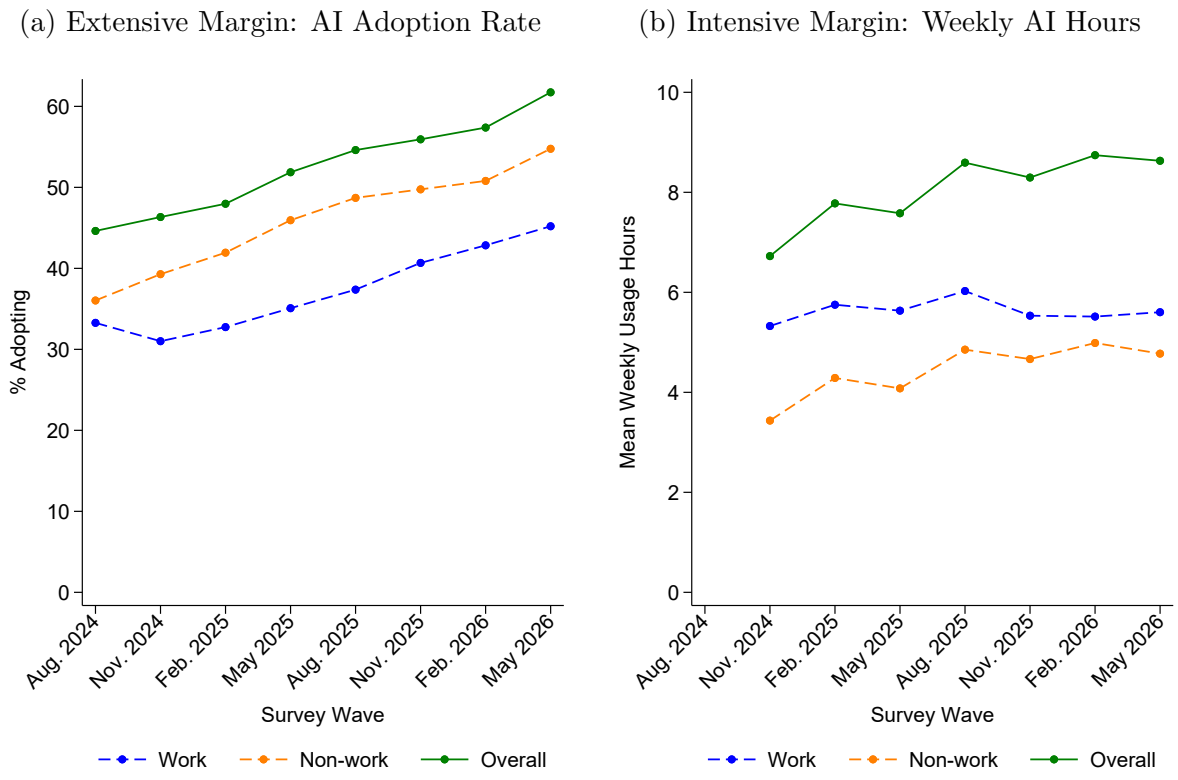
> *Generative AI is a type of artificial intelligence that creates text, images, audio, or video in response to prompts. Some examples of Generative AI include ChatGPT, Gemini, and Midjourney.*

We included examples of popular genAI products because we thought some respondents may be more familiar with those product names than with the broader concept of genAI. We then ask respondents:

> Employed Respondents: *Do you use Generative AI for your job? (No/Yes)*
>
> All Respondents: *Do you use Generative AI [outside your job]? (No/Yes)*

[3] The few observations above the 45 degree line reflect modest differences in employment shares between the RPS and CPS at the 3-digit level; see Appendix Figure D.3a. Appendix Figure D.3b is the equivalent of Figure 3a using the RPS both for the x- and y-axis; it shows quantitatively very similar results, except that no observations lie above the 45 degree line.

Figure 4: GenAI Adoption Over Time



Notes: Panel 4a shows the share of respondents using genAI for work, non-work, and overall by RPS survey wave. The sample for the "Work" series is employed respondents. Panel 4b shows the mean usage hours per person by survey wave, conditional on using genAI for work, non-work, or overall. Information on time spent using genAI is not comparable to the other survey waves in the Aug. 2024 survey, so it is excluded from Panel 4b. See Bick et al. (2026b) for how the RPS elicits the information to construct weekly usage hours.

These questions are designed to mirror analogous questions measuring computer adoption from the CPS Computer and Internet Use Supplement (CIU), which were asked between 1984 - 2003; see Appendix Section A.5. For the second question, non-employed respondents are not shown the term in brackets. The survey asks several follow-up questions of genAI users, including which products they used, how intensively they used it, and estimates of how much time it saved them. See Bick et al. (2026b) for an overview of results.

A caveat is that our measurement approach will only capture genAI use that respondents are aware of. For example, while typing prompts into ChatGPT is an obvious case of using genAI, automatic AI-assisted writing suggestions are a more subtle case that respondents may not consider. Because our estimates may not detect some passive or embedded uses of the technology, we interpret our estimates as a lower bound on the extent of genAI adoption.

Figure 4a plots the share of individuals who report using genAI for work, non-work, and overall. As of May 2026, 55% of US adults aged 18-64 use genAI for non-work reasons, while 45% of workers use genAI for work. The overall adoption rate, i.e., the share of individuals who use genAI either for work or non-work purposes, is 62%. Along all three margins, adoption has increased since August 2024.

Figure 4b shows the mean hours spent using genAI, among users, in the past week, providing an indication of the intensity of genAI use over time. (The series begin in November 2024 because our time use categories changed between the August and November 2024 waves.) Mean weekly hours among genAI users are flat for work users, but increase mildly for non-work use.

2.5 Measuring Task-Level Generative AI Use

We can construct an occupation-level measure of genAI adoption, which indicates the share of people in a CPS detailed (3-digit) occupation who report using genAI for work. Figure 5a shows that occupation-level genAI adoption rates are strongly correlated across four waves of the RPS, from August 2025 to May 2026.

GenAI users are asked which of their work tasks genAI regularly helps them perform. This allows us to construct a task-level measure of genAI adoption, which is the share of workers who use genAI for a given task (among all workers who perform that task). Figure 5b shows that task-level genAI adoption rates are strongly correlated across four waves of the RPS.

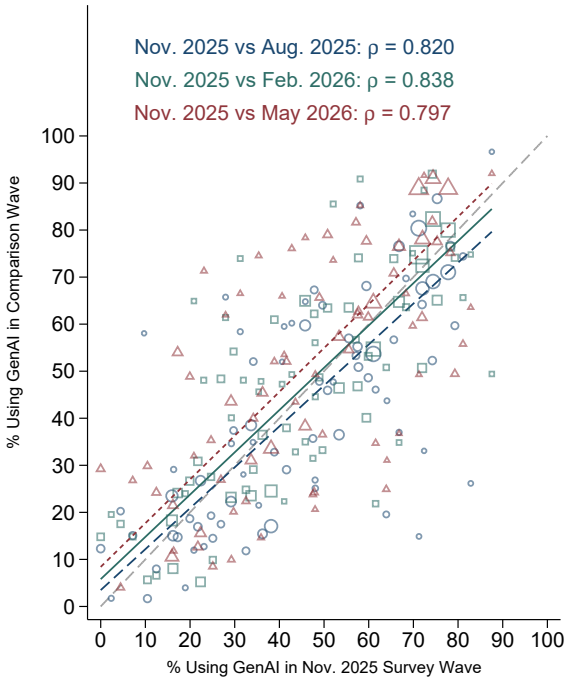
The stability of occupation- and task-level adoption rankings across survey waves provides evidence that our measures of genAI adoption are internally consistent across RPS survey waves and helps address concerns over measurement error.

3 How Widespread is Generative AI Adoption?

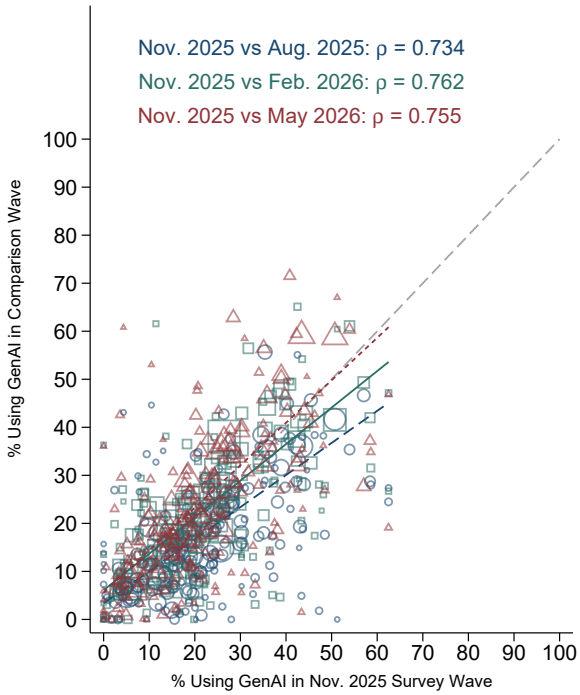
This section analyzes genAI adoption rates across occupations and tasks using the RPS. To guide the analysis, we introduce a very simple model of production that clarifies how RPS estimates map into underlying economic objects of interest. The model is deliberately stylized, but could be extended to incorporate additional channels through which genAI affects economic outcomes. In later sections, the model serves as a common lens through which we can interpret and compare alternative genAI measures.

Figure 5: Consistency of GenAI Adoption Rates Across Survey Waves

(a) Occupation-Level



(b) Task-Level



Notes: In Panel (a), an observation is an occupation at the CPS detailed (3-digit) level. In Panel (b), an observation is an ONET Intermediate Work Activity. Circle sizes correspond to observation weights. We include only observations with at least 10 observations in each survey wave. The number reported as ρ is the weighted correlation coefficient between adoption weighted by number of appearances of the occupation/task in the pooled data.

3.1 Conceptual Framework

Economic activity consists of a set of tasks, indexed by $t \in \{1, \ldots, T\}$. Workers are assigned to occupations $o \in \{1, \ldots, O\}$, with employment shares λ_o, where $\sum_o \lambda_o = 1$. In occupation o, task t's share of work is $\omega_{o,t} \geq 0$, where $\sum_t \omega_{o,t} = 1$. Aggregate output is a linear aggregation of occupation-task production:

$$Y = \sum_o \lambda_o \sum_t \omega_{o,t} y_{o,t} \tag{1}$$

where $y_{o,t}$ denotes task-level output, assumed to be one prior to the introduction of genAI.

We model genAI as a technology that increases workers' productivity for a subset of tasks. Among workers who perform task t, a share $a_t \in [0, 1]$ adopts genAI for that task.

Conditional on adoption, genAI increases task productivity by a factor g_t. Aggregate output after the introduction of genAI is therefore given by

$$Y' = \sum_o \lambda_o \sum_t \omega_{o,t}(1 + a_t g_t), \tag{2}$$

Given $y_{o,t} = 1$, the percent change in aggregate productivity from genAI is

$$\frac{Y' - Y}{Y} = \sum_t a_t g_t \sum_o \lambda_o \omega_{o,t} = \sum_t \omega_t a_t g_t \tag{3}$$

where $\omega_t \equiv \sum_o \lambda_o \omega_{o,t}$ is the total share of task t in the economy.

Within this framework, the economic impact of genAI depends on three key objects. The first is ω_t, the economy-wide share of task t. The RPS provides direct estimates of ω_t by measuring the share of workers who report performing task t (see Figure 2b). Two limitations of these estimates are worth noting. First, while they capture the extensive margin of task performance (i.e., the share of workers who perform a task), they may not fully capture differences in task intensity (e.g., total share of work time spent on a task). Second, the RPS omits tasks outside the top ten for each occupation; however, as shown in Figure 3b, this restriction does not appear to bias estimates of a task's prevalence in the economy.

The second object is task-level genAI adoption rates, a_t. Adoption rates determine the extent to which potential productivity gains can be realized in practice. The RPS provides estimates of a_t: among workers who report performing task t, we observe the share of workers who report using genAI for that task. We summarize these estimates below in Section 3.2.

The final object is task-level productivity gains from genAI conditional on adoption, g_t. The RPS contains self-reported estimates of genAI time savings, which could be combined with information on task adoption rates to estimate g_t. Complementary evidence on g_t can be drawn from micro-level experimental studies for selected tasks or from exposure scores.

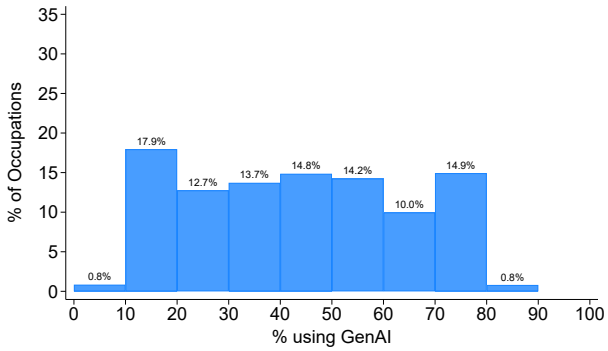
In summary, subject to the limitations discussed above, the RPS provides direct estimates or empirical proxies for each of the objects needed to bring the framework to the data.

3.2 Distribution of Generative AI Adoption Rates Across Occupations and Tasks

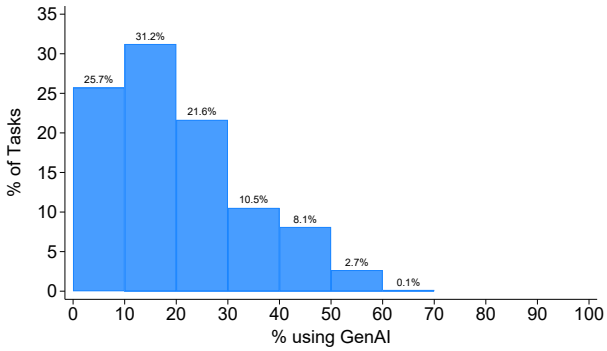
Figure 6 displays the distribution of adoption rates across detailed occupations and tasks. Together, these distributions point to three key takeaways. First, genAI adoption is already

Figure 6: Distribution of GenAI Adoption Rates by Occupation and Task

(a) GenAI Adoption Rates by Occ. (3-digit SOC)



(b) GenAI Adoption Rates by Task (DWA)



Notes: Tasks are at the Detailed Work Activity (DWA) level. Figure includes only occupations (tasks) with at least 20 observations in the pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 RPS survey waves. Tasks are weighted by the number of observations in the pooled survey waves, and occupations are weighted by the number of people. Occupation is at the CPS detailed (3-digit) occupation level.

widespread and assists a broad range of work. For example, over 80% of occupations and 40% of tasks exhibit adoption rates exceeding 20%.

Second, even within detailed occupations and tasks, adoption outcomes are often highly variable across individuals. For example, 40% of occupations have adoption rates between 20% and 50%. This implies that in a large share of occupations, at least a fifth of workers use genAI, while more than half do not. Similarly, roughly 40% of tasks have adoption rates between 20% and 50%. The present distribution of genAI adoption can therefore be characterized as widespread but shallow: genAI is used in a large share of detailed occupations and tasks, but in most occupations and tasks adoption is far from universal. This variability within detailed occupations and tasks suggests important individual-level determinants of adoption, a point we explore further in Section 5.

A final takeaway is that, although most occupations exhibit low-to-moderate adoption rates, a few occupations do exhibit very high adoption rates. For example, roughly 15% of occupations exhibit adoption rates above 70%. By contrast, only 2.8% of tasks have adoption rates above 50% and none exceed 70%. This implies that high-adoption occupations are not driven by the presence of a single task with near-universal adoption. Rather, high-adoption occupations are those with a collection of moderately-high-adoption tasks, which together imply that a large majority of workers in the occupation uses genAI for at least one part of their job.

Table 1: Occupations and Tasks With Highest and Lowest genAI Adoption Rates

Top 10 occupations		**Top 10 tasks**	
Occupation	**Rate (%)**	**Task**	**Rate (%)**
Computer and information research scientists	87.3	Read documents to gather technical information.	61.3
Information security analysts	85.4	Prepare research reports.	60.7
Network and computer systems administrators	82.4	Design integrated computer systems.	59.5
Computer programmers	81.2	Analyze data to identify trends or relationships among variables.	57.5
Public relations specialists	80.4	Update knowledge about emerging industry or technology trends.	57.1
Personal financial advisors	79.7	Conduct quantitative failure analyses of operational data.	55.8
Chief executives	79.7	Prepare data for analysis.	55.4
Software developers	78.8	Develop marketing plans or strategies.	55.0
Market research analysts and marketing specialists	78.7	Collaborate with others to determine design specifications or details.	54.8
Computer and information systems managers	78.2	Design websites or web applications.	54.6
Bottom 10 occupations		**Bottom 10 tasks**	
Occupation	**Rate (%)**	**Task**	**Rate (%)**
Stockers and order fillers	14.8	Arrange delivery of goods or services.	0.0
Janitors and building cleaners	13.8	Assist practitioners to perform medical procedures.	0.0
Other installation, maintenance, and repair workers	13.1	Collect biological specimens from patients.	0.0
Maids and housekeeping cleaners	11.3	Drive trucks or other vehicles to or at work sites.	0.0
Fast food and counter workers	11.3	Inspect electrical or electronic systems for defects.	0.0
Maintenance and repair workers, general	11.2	Maintain medical equipment or instruments.	0.0
Painters and paperhangers	10.4	Measure the physical or physiological attributes of patients.	0.0
Licensed practical and licensed vocational nurses	10.4	Prepare patient treatment areas for use.	0.0
Receptionists and information clerks	7.6	Present food or beverage information or menus to customers.	0.0
Animal caretakers	5.3	Record service or repair activities.	0.0

Notes: Occupations are defined using CPS detailed (3-digit) occupation codes. Tasks are Detailed Work Activities (DWA). "Rate" for occupations is the share of workers in that occupation who report using genAI at work (pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 RPS survey waves). "Rate" for tasks is the share of worker-task observations in which the worker reports using genAI for that task. The occupation (task) panel excludes occupations (tasks) with fewer than 20 observations in the pooled waves.

3.3 High- and Low-Adoption Occupations and Tasks

The left panel of Table 1 shows the occupations with the highest and lowest adoption rates. The highest-adoption occupations are concentrated in computing and professional roles, with adoption rates above 75%. These occupations are characterized by work that involves generating, processing, or communicating information. In contrast, the lowest-adoption occupations are primarily in personal services, transportation, and manual labor, with adoption rates below 15%. These occupations tend to involve in-person interaction or physical tasks.

The right panel of Table 1 shows the tasks (DWAs) with the highest and lowest adoption rates. High-adoption tasks involve organizational, analytical, and information-processing activities, such as reading and preparing reports, developing models, and analyzing data, with adoption rates between 55–61%. In contrast, low-adoption tasks involve physical, operational, or situational activities, such as working with patients or work equipment, driving vehicles, or landscaping work. Appendix Figure D.2 aggregates adoption rates from all tasks into 37 Work Activity adoption rates. Similar patterns emerge in this figure, with low adoption rates in physical and personal care activities and high adoption rates in computer- and analysis-related activities. A full tabulation of adoption rates by detailed occupation and task can be downloaded from https://sites.google.com/view/covid-rps/generative-ai.

4 Does Generative AI Exposure Predict Adoption?

Several papers construct predictions of AI "exposure," which attempt to map AI capabilities into the task or occupation space (Brynjolfsson et al., 2018; Eloundou et al., 2024; Felten et al., 2021; Webb, 2020). The goal is to predict which tasks or occupations could be most directly affected by AI. Before presenting the empirical comparisons, we first show how several widely-used exposure scores map into the conceptual framework laid out in Section 3.1.

4.1 Mapping Exposure Scores Into the Conceptual Framework

Eloundou et al. (2024) construct task-level exposure scores based on whether they think LLM's with ChatGPT 4.0 capabilities could (at least) double productivity in a given task. We can represent such an exposure score e as an indicator function which equals one if a

task's predicted productivity gain, $\tilde{g}_t^e$, exceeds a threshold $\overline{g}^e$:

$$e_t = \mathbb{1}\{\tilde{g}_t \geq \overline{g}^e\} \tag{4}$$

In the case of Eloundou et al. (2024), $\overline{g}^e = 1$, i.e., the threshold is a 100% productivity gain.

Eloundou et al. (2024) also construct occupation-level exposure scores by computing the share of an occupation's ONET tasks that are classified as exposed. Felten et al. (2021) predict occupational exposure by linking the skill requirements of occupations to the demonstrated capabilities of AI systems. Occupations that depend more on abilities that AI can already replicate are considered more exposed. Brynjolfsson et al. (2018) predict occupation-level exposure to machine learning. Webb (2020) creates a patent-based exposure measure, which predicts the extent to which occupations are exposed to automation based on the overlap between job tasks and patented innovations. Because these latter three measures are not constructed at the task level they do not map directly into our framework, though intuitively they play similar roles as proxies for task-level productivity gains aggregated to the occupation.

4.1.1 A Simple Extension: From Exposure to Adoption

A simple extension of our framework introduces an adoption decision by workers, which provides a natural link between exposure scores and observed genAI use. Assume that worker i faces an idiosyncratic adoption cost $\chi_{i,t}$ if they adopt genAI for task t. Workers adopt genAI for task t only if the associated productivity gain exceeds this cost, i.e., if $g_t > \chi_{i,t}$. Let F_t denote the distribution of genAI adoption costs among workers who perform task t in their job. Then the task-level adoption rate is

$$a_t = \Pr(g_t \geq \chi_{i,t} | i \text{ performs } t) = F_t(g_t). \tag{5}$$

The equation implies that a task t that is exposed ($e_t = 1$) will have a higher adoption rate than another task s that is not exposed ($e_s = 0$) if the two tasks have the same adoption cost distribution, $F_t = F_s$.

Adoption costs can be viewed as a combination of task-specific, worker-specific, and idiosyncratic barriers to adoption:

$$\chi_{i,t} = \mu_t + \xi_i + \epsilon_{i,t} \tag{6}$$

Task-specific barriers μ_t reflect barriers that are common to all workers. For example, some tasks may require costly experimentation before genAI can be effectively leveraged, or may entail the use of sensitive information in which genAI use is discouraged. Worker-specific barriers ξ_i capture persistent differences in individuals' propensity to adopt, which may reflect heterogeneous skills, beliefs, preferences, or prior experience with genAI. The residual term $\epsilon_{i,t}$ captures barriers to adoption that vary within a worker across tasks. For example, some workers may find it easier to use genAI to help with coding than to help with brainstorming, while a different worker may have the reverse ranking.[4]

This framework highlights three reasons why exposure scores may not perfectly correlate with adoption rates. First, some exposed tasks may also have higher average task-specific adoption costs, μ_t. This could be because, for example, the task involves confidential data that managers do not want fed into a genAI platform. Second, some exposed tasks could have a high average ξ_i among workers who perform that task, perhaps because the workers tend to lack complementary computer or analytical skills. Finally, exposure predictions may be inaccurate; that is, they may not perfectly predict genAI's productivity gain at the task level, g_t.

The remainder of this section compares exposure predictions with adoption rates in the RPS. We then highlight occupations and tasks whose adoption rates diverge from predicted exposure and discuss potential sources of adoption barriers.

4.2 Do Exposure Scores Predict Adoption?

4.2.1 Occupation-Level Exposure Scores

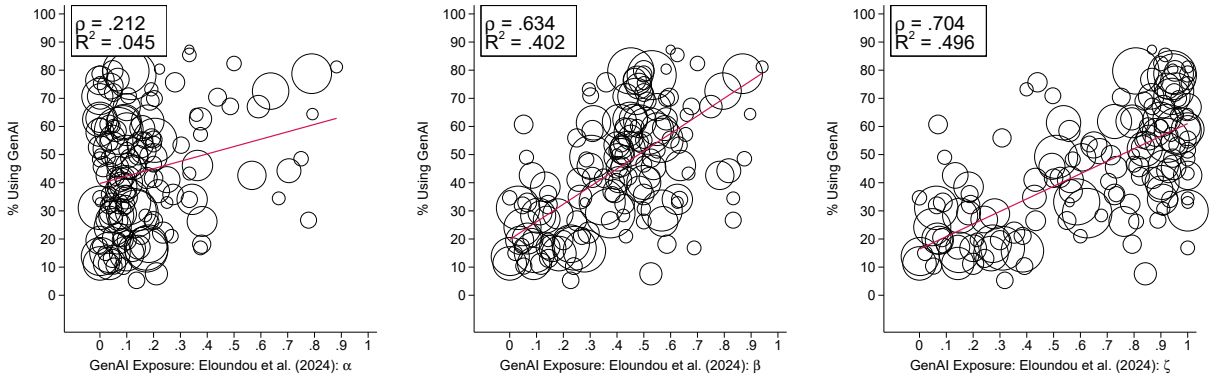
Figure 7 compares occupation-level genAI adoption rates from the RPS to six genAI exposure scores. Figures 7a–7c compare adoption to three measures from Eloundou et al. (2024), α, β, and ζ, which assume progressively greater genAI capabilities.[5] The α measure, which assumes the lowest level of genAI capabilities among the Eloundou et al. (2024) measures, and the measure of Brynjolfsson et al. (2018), which was constructed several years earlier than the other measures, are only weakly correlated with actual adoption. The β and ζ measures, which assume greater capabilities, are more strongly correlated with adoption, as

[4]Lindenlaub et al. (2026) develop a related structural framework in which genAI adoption is based on comparative advantage and may therefore differ from predicted exposure.

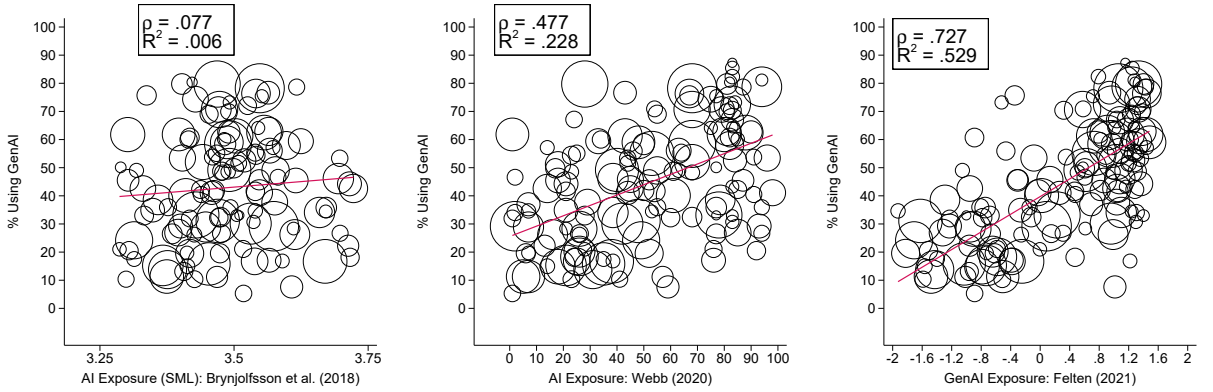
[5]We use the original GPT-4-based scores throughout. Yin et al. (2026) show that the Eloundou et al. (2024) methodology produces notably different exposure scores when applied to more recent models. Bick et al. (2026b) and Lindenlaub et al. (2026) also compare adoption rates and exposure scores by occupation.

Figure 7: GenAI Adoption versus Exposure Predictions by Occupation

(a) Eloundou et al. (2024): α (b) Eloundou et al. (2024): β (c) Eloundou et al. (2024): ζ



(d) Brynjolfsson et al. (2018) (e) Webb (2020) (f) Felten et al. (2021)



Notes: Each observation in this regression is a the CPS detailed occupation. Circle sizes correspond to regression weights (size of the occupation in people in the pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 RPS survey waves). Figure includes only occupations with at least 20 observations in the pooled data. The number inside the box, ρ, is the weighted correlation coefficient between occupation-level adoption and the given exposure measure.

are the measures from Webb (2020) and Felten et al. (2021).

The positive correlations between exposure predictions and realized adoption help validate that these widely used measures capture meaningful variation in which occupations are assisted by genAI. At the same time, all correlations remain well below one, indicating gaps between predicted exposure and actual adoption to date.

Figure 7 evaluates exposure scores on their ability to rank occupation adoption rates. Given the substantial within-occupation variation in adoption documented in Section 3.2, a natural next question is how well occupation-level measures predict adoption for a particular worker. Column 1 of Table 2 reports this comparison for three such measures.

Table 2: Predicting GenAI Adoption at the Worker and Task Levels

Controls	Adjusted R^2	
	Worker Level	Task Level
Detailed Occupation Fixed Effects	0.176	0.111
DWA Fixed Effects		0.107
Eloundou et al. (2024) ζ Exposure Measure	0.081	0.028
Felten et al. (2021) Exposure Measure	0.086	

Notes: Observations are at the worker or worker-task level depending on the column. A task is a detailed work activity (DWA). The dependent variable in all regressions is an indicator for whether genAI is used at work. Uses pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 survey waves, and all regressions include a fixed effect for survey wave.

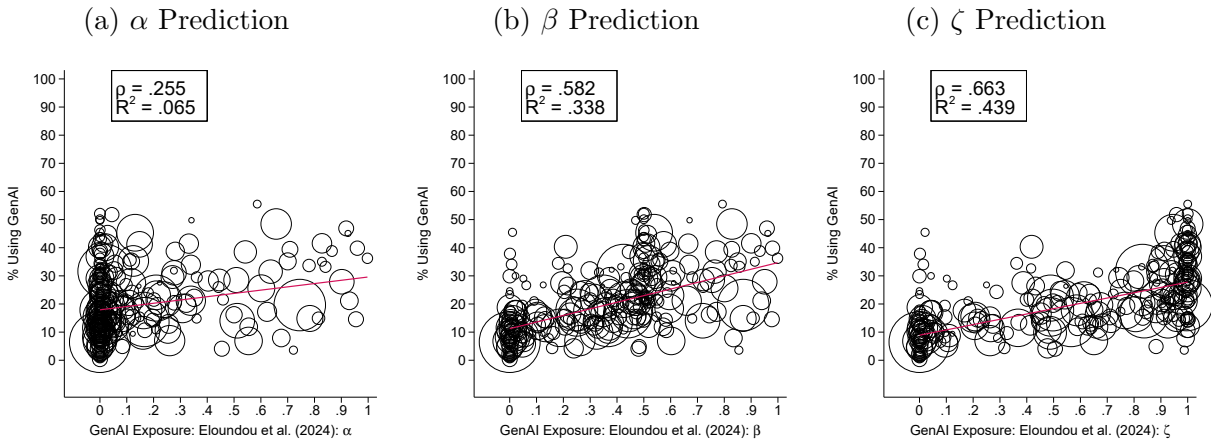
The top row shows that occupation fixed effects have an adjusted R^2 of 0.176, implying that less than 20% of the variation in worker-level adoption is explained by occupation. The next two rows report results for two occupation exposure scores: the ζ exposure measure from Eloundou et al. (2024) and the Felten et al. (2021) exposure measure. These exposure scores yield adjusted R^2 of 0.081 and 0.086, respectively. That is, exposure scores capture slightly less than half of the occupation-level variation in adoption that is available to be explained.

To summarize, we find that occupation exposure scores are predictive of actual genAI adoption, though they explain only about half of the variation in adoption across occupations. Moreover, because occupation itself explains less than 20% of variation in worker-level adoption, occupational exposure scores are only mildly predictive of worker-level adoption.

4.2.2 Task-Level Exposure Scores

Figure 8 compares task-level adoption rates in the RPS to the three task-level exposure scores from Eloundou et al. (2024). The pattern mirrors the occupation-level results: the α measure is only weakly correlated with adoption, while the β and ζ measures, which assume greater genAI capabilities, are more strongly correlated. Exposure scores are therefore also predictive of which tasks genAI currently assists. However, the correlations between adoption and the β and ζ exposure scores are somewhat lower at the task level than at the occupation level, indicating larger gaps between task-level exposure predictions and realized adoption.

Figure 8: Eloundou et al. (2024) Exposure Predictions versus Adoption by Work Task



Notes: Each observation in this regression is a task, with tasks at the Intermediate Work Activity (IWA) level. Circle sizes correspond to regression weights (number of appearances of the task in the pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 RPS survey waves). Figure includes only tasks with at least 20 observations in the pooled data. The number inside the box, ρ, is the weighted correlation coefficient between task-level adoption and the given Eloundou et al. (2024) task measure.

Column 2 of Table 2 repeats the exercise at the worker-task level, where the dependent variable indicates whether the worker uses genAI for a given task. The top row shows that DWA fixed effects have an adjusted R^2 of 0.107, implying that tasks explain just over 10% of adoption variation at the worker-task level. The ζ exposure measure yields an adjusted R^2 of 0.028, less than a third of what tasks themselves explain.

Interestingly, detailed occupation fixed effects explain a slightly larger share of variation in worker-task adoption than task fixed effects, 0.111 versus 0.107, respectively. That is, a worker's occupation is more informative about whether the worker uses genAI for a particular task than the task itself. One potential explanation is that some determinants of adoption operate at the worker level rather than the task level, so that experience with genAI in one part of a job carries over to others. We present evidence consistent with this in Section 5.

To summarize, exposure scores predict task-level adoption, but explain less than a third of the variation that task fixed effects capture, an even smaller share than at the occupation level.

Table 3: Occupations and Tasks Most Over- and Under-Predicted by Exposure-Based Measures

Panel A: Occupations

Most over-predicted occupations				**Most under-predicted occupations**			
Occupation	Act.	Pred.	Gap	Occupation	Act.	Pred.	Gap
Receptionists and information clerks	7.6	54.0	-46.4	Construction equipment operators	60.7	19.5	41.2
Medical secretaries and administrative assistants	16.8	61.0	-44.2	Computer and office machine repairers	75.7	36.0	39.7
Tellers	18.1	51.8	-33.6	Industrial electrical and electronics repairers	73.2	34.2	39.0
Customer service representatives	30.1	61.0	-30.9	Special education teachers	70.9	38.6	32.3
Data entry keyers	26.7	56.1	-29.4	Computer and information research scientists	87.3	55.1	32.2
Operations research analysts	33.0	61.0	-28.0	Laundry and dry-cleaning workers	49.0	20.6	28.5
Insurance claims and policy processing clerks	34.4	61.0	-26.6	Information security analysts	85.4	57.3	28.1
Logisticians	30.7	57.0	-26.2	Chief executives	79.7	52.4	27.3
Food preparation and serving supervisors	26.2	52.4	-26.2	Network and computer systems administrators	82.4	56.6	25.8
Animal caretakers	5.3	30.6	-25.3	Shuttle drivers and chauffeurs	55.9	30.1	25.8

Panel B: Tasks

Most over-predicted tasks				**Most under-predicted tasks**			
Task	Act.	Pred.	Gap	Task	Act.	Pred.	Gap
Operate communications equipment or systems.	3.6	27.5	-23.9	Obtain formal documentation or authorization.	45.5	9.3	36.2
Prescribe medical treatments or devices.	4.9	25.5	-20.7	Test sites or materials for environmental hazards.	37.9	8.9	28.9
Check materials/documents for accuracy or compliance.	7.9	27.6	-19.7	Develop models of systems, processes, or products.	55.5	27.8	27.7
Estimate project development or operational costs.	10.7	27.8	-17.1	Position tools or equipment.	34.1	9.2	24.9
Assist others to access additional services or resources.	12.3	27.8	-15.5	Analyze performance of systems or equipment.	51.8	27.0	24.8
Operate office equipment.	4.1	18.4	-14.3	Develop business or marketing plans.	52.1	27.8	24.3
Notify others of emergencies or problems.	6.9	21.2	-14.3	Resolve personnel or operational problems.	40.3	16.8	23.5
Evaluate patient/client conditions or treatments.	4.2	17.9	-13.7	Investigate organizational or operational problems.	49.7	26.6	23.1
Order medical tests or procedures.	14.3	27.8	-13.5	Research laws, precedents, or other legal data.	50.1	27.8	22.4
Prepare health or medical documents.	14.6	27.8	-13.2	Evaluate project feasibility.	49.8	27.8	22.0

Notes: Panel A reports occupations with the largest negative (over-predicted) and positive (under-predicted) residuals from a regression of occupation-level genAI adoption rates on exposure scores constructed using Eloundou et al. (2024). Occupation is at the CPS detailed (3-digit) occupation level. Panel B reports the analogous task-level results using ONET tasks. Observed adoption is the actual genAI adoption rate, predicted adoption is the fitted value based on exposure, and the residual is observed minus predicted adoption. Rows are ordered by the absolute value of the residual. All statistics are based on pooled Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 survey waves.

4.3 Where Do Exposure and Adoption Diverge?

Which occupations and tasks exhibit large gaps between adoption and predicted exposure? For simplicity, our comparisons focus on the ζ exposure scores from Eloundou et al. (2024), which are available at both the occupation and task level. We regress RPS adoption rates on these scores, which yields a predicted adoption rate for each occupation and task. Table 3 summarizes the occupations and tasks whose actual adoption rates deviate most from this prediction, either below ("over-predicted") or above ("under-predicted").

Many over-predicted occupations are clerical and administrative. For example, predicted adoption exceeds actual adoption by 46 percentage points for receptionists and information clerks, 44 points for medical secretaries and administrative assistants, 34 points for tellers, and 31 points for customer service representatives. Data entry keyers and insurance claims and policy processing clerks are also over-predicted. Within our framework, two channels plausibly contribute to this pattern. First, worker-level adoption costs, ξ_i, may be higher on average for these workers, potentially reflecting limited complementary technology skills or weak incentives to adopt. Second, task-level frictions, μ_t, may be high in these roles: standardized workflows, compliance requirements, and restrictions on the use of confidential data may prevent widespread adoption of genAI.

The most over-predicted tasks are consistent with these explanations. They include examining materials or documentation for accuracy or compliance, prescribing medical treatments or devices, evaluating patient or client conditions, ordering medical tests or procedures, and preparing health or medical documents. Regulatory, privacy, and liability concerns may constrain genAI use in these tasks. Other over-predicted tasks include operating communications or office equipment and estimating project development or operational costs. Although genAI could potentially assist these tasks, capturing those benefits may require complementary capital investments, access to organizational data, or costly workplace reorganization.

Under-predicted occupations include a variety of occupation types, including construction equipment operators, several types of computer and electronic-equipment repairers, special education teachers, computer and information research scientists, information security analysts, chief executives, and network and computer systems administrators. These workers tend to have a relatively high degree of autonomy in their jobs, which may reflect low values of ξ_i.

These workers may also face low task-level frictions, μ_t, in their jobs: coding, lesson plan design, and directions are all tasks that genAI can assist with minimal AI-specific train-

Table 4: Predicting Adoption: Individuals vs. Tasks

Controls	Adjusted R^2
DWA Fixed Effects	0.112
DWA Fixed Effects + Demographic Controls	0.127
DWA Fixed Effects + Individual Fixed Effects	0.440

Notes: Tasks are at the Detailed Work Activity (DWA) level. Observations are at the worker-task level, and a given observation is whether the worker who is doing that task reports using genAI for it. Includes only people who do at least two tasks. Each regression has 55,634 person-task observations. Demographic Controls consist of education, race, sex, and a quadratic for age.

ing. Under-predicted tasks include developing models of systems or processes, analyzing the performance of systems or equipment, developing business or marketing plans, investigating organizational problems, researching legal information, and evaluating project feasibility, which genAI appears well-suited to assist given the requisite knowledge of what needs to be done.

5 Learning and the Diffusion of GenAI Across Domains

While genAI adoption rates vary across different tasks, adoption is also highly variable among individuals performing the same task. For example, Figure 6 showed that for roughly 40% of tasks, at least a fifth of workers use genAI but more than half do not. Why do some workers adopt genAI for a given task while others do not?

One potential explanation for within-task variation in adoption is that some workers are systematically more likely to adopt genAI than others. Table 4 tests this by regressing worker-task-level genAI adoption on task fixed effects and individual-level controls. Task fixed effects alone yield an adjusted R^2 of only 0.112.[6] This means that tasks themselves account for roughly 11% of the variation in adoption, leaving the vast majority unexplained.

The third row of Table 4 shows that including individual fixed effects dramatically increases the adjusted R^2 to 0.440. This sharp increase implies that a key driver of adoption is

[6]Table 2 reports a slightly lower adjusted R^2 of 0.107 for the same regression. This reflects slightly different samples: in Table 4, we only include individuals who perform at least two tasks to keep the sample consistent across all specifications in that table.

that some workers are consistently more likely to use genAI than others, regardless of task.

What drives this persistent individual variation? The second row of Table 4 shows that controlling for demographic characteristics only marginally improves predictive power, indicating that most individual differences occur within demographic groups rather than between them. This suggests that individual characteristics, such as experience, skills, beliefs, willingness to experiment, or the individual's work environment, play a central role in adoption decisions.

Our dataset does not allow us to investigate most of these characteristics directly. What it does contain, however, is information on workers' adoption choices across multiple domains: in the past and the present, across tasks, and at work versus at home. This allows us to investigate one specific channel that could generate persistent individual differences: learning from genAI experience that carries over across domains. Using genAI in one domain may teach workers about genAI's strengths and limitations, how to effectively prompt, and how to integrate it into workflows, reducing the need to learn or experiment again in other domains.

If using genAI in one domain lowers the cost of adopting it in another, workers who use genAI in one domain should also use it more in others. We begin by formalizing this idea within our conceptual framework and then proceed with an empirical investigation.

5.1 Conceptual Framework: Learning as a Cost Shifter

Recall from Section 4.1.1 that worker i adopts genAI for task t when the productivity gain exceeds the adoption cost:

$$a_{i,t} = \mathbf{1}\{g_t \geq \chi_{i,t} = \mu_t + \xi_i + \varepsilon_{i,t}\}. \tag{7}$$

The worker component ξ_i captures everything that makes a given individual systematically more or less likely to adopt across tasks. The fixed-effect results above indicate that ξ_i accounts for a large share of adoption variation.

To formalize a notion of learning, we let ξ_i itself depend on the worker's accumulated genAI experience:

$$\xi_i = \bar{\xi}_i - \theta \cdot E_i, \tag{8}$$

where $\bar{\xi}_i$ is an initial underlying individual propensity to adopt, E_i measures the worker's experience with genAI, and $\theta \geq 0$ governs the impact of learning on the fixed cost of adoption. Under this specification, experience lowers the cost of adoption, so workers with larger E_i

are more likely to adopt on any given task.

Our goal is to assess whether the learning term $\theta \cdot E_i$ is a quantitatively meaningful component of ξ_i. The empirical challenge is that $\bar{\xi}_i$ may be negatively correlated with E_i: workers with low $\bar{\xi}_i$ are more likely to adopt genAI and therefore accumulate experience. This would imply that a raw correlation between adoption and proxies for experience will reflect a combination of learning and selection. The remainder of this section uses three sources of variation in E_i to isolate the learning component: one based on a retrospective question of past genAI use, and two based on concurrent genAI use in other domains. The three tests are complementary. The retrospective test speaks most directly to experience accumulated over time. The two concurrent tests speak to whether experience is portable across domains. No single test is decisive, but similar findings across distinct analyses makes the learning interpretation more credible than any single analysis on its own.

5.2 Three Empirical Tests of the Learning Channel

Past GenAI Use and the Breadth of Current Adoption Our first test exploits variation in the length of time that workers have used genAI. Our dataset asks whether workers were already using genAI six months before the survey, which we use as a binary proxy for E_i. If workers incur learning or setup costs when they first adopt genAI, but not when extending its use to additional tasks, then more experienced users should use it for more tasks. The identifying assumption is that, among current genAI users, genAI experience is uncorrelated with the worker's underlying propensity to adopt, $\bar{\xi}_i$, after conditioning on occupation, demographics, and the total number of tasks performed.

Table 5 provides evidence consistent with this prediction. Among workers who report using genAI for at least one task, those who have used genAI for six months or more report using it for 0.4 more tasks currently. On average, workers report performing 4.7 tasks, so this corresponds to a 9% increase.

The main threat to this interpretation is that workers who were already using genAI six months ago plausibly have lower $\bar{\xi}_i$ than workers who only adopted more recently. If so, part of the coefficient reflects selection among earlier adopters rather than learning. The next two tests address this concern using sources of variation that are less directly tied to a worker's underlying propensity to adopt.

Table 5: Prior GenAI Experience and Breadth of Task-Level Adoption

	Number of genAI-Assisted Tasks
Have used genAI 6 months ago	0.429***
	(0.113)
Adjusted R^2	0.441
Observations	6,565

Notes: Regression is at the worker level and includes only workers who report using genAI for at least one task. The dependent variable is the number of genAI-assisted tasks the worker performs. The key independent variable is an indicator for whether the worker reports having used genAI for at least six months. The regression controls for a quadratic in age, detailed occupation, survey wave fixed effects, sex, and the total number of tasks performed by the worker. Regression includes a constant term that is not in the table. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Cross-task spillovers Our second test asks whether workers are more likely to use genAI in a given task if they use genAI in other tasks. We begin by constructing, for each worker-task observation, the share of the individual's other tasks for which they use genAI. Because the worker's genAI use in other tasks is likely endogenous to $\bar{\xi}_i$, we then construct an Other-Task Adoption Index (OAI): for each of a worker's other tasks, we compute the adoption rate among all other workers performing that task, and then average these leave-one-out rates across the worker's other tasks. Intuitively, the OAI captures the worker's expected adoption rate for their other tasks if they behaved like the average worker performing those tasks. The OAI serves as an instrument for the worker's behavior in other tasks that is not directly driven by $\bar{\xi}_i$, the worker's underlying propensity to adopt. We then run a 2SLS regression of adoption on the focal task using the OAI instrument and include task fixed effects, demographics, and occupation fixed effects.

The identifying assumption is that, within a given occupation, the composition of a worker's task portfolio is uncorrelated with the worker's adoption propensity, $\bar{\xi}_i$. This rules out, for example, low-$\bar{\xi}_i$ workers sorting into task portfolios with systematically higher adoption rates within the same occupation.

The strength of this assumption depends on the detail of the occupation controls used. If we eliminate occupation controls, we are assuming that workers do not sort across occupations according to their underlying adoption propensity, $\bar{\xi}_i$, which is unlikely to hold in practice. More detailed occupation controls weaken this requirement, but they also absorb

Table 6: Spillovers from High-Adoption Tasks to Other Tasks

	(1)	(2)
Broad occupation controls	No	Yes
Panel A. OLS: Dependent variable = Uses genAI for focal task		
Other-task AI use share	0.628***	0.621***
	(0.011)	(0.011)
Observations	55091	55091
Panel B. 2SLS: Other-task AI use share instrumented by OAI		
Other-task AI use share	0.702***	0.545***
	(0.039)	(0.067)
First-stage F-stat.	436.37	92.97
Observations	55091	55091
Panel C. First stage: Dependent variable = Other-task AI use share		
Other-task adoption index (OAI)	0.818***	0.615***
	(0.039)	(0.064)
Observations	55091	55091

Notes: The unit of observation is a worker-IWA pair. The dependent variable in Panels A and B is an indicator for whether the worker uses genAI for the focal IWA. Other-task AI use share is the share of the worker's other reported IWAs for which the worker uses genAI. The instrument, Other-task adoption index (OAI), is the leave-one-out mean IWA-level adoption rate of the worker's other IWAs, excluding the worker's own adoption decisions. To construct IWA-level outcomes, the data are first collapsed to unique worker-IWA observations, with IWA-level genAI use equal to one if the worker uses genAI for any DWA within that IWA. Analysis includes only IWAs with at least 20 observations. All regressions include controls for education, survey wave, race, sex, age, age squared, the total number of IWAs performed by the worker, and IWA fixed effects. Columns differ in whether broad occupation controls are omitted or included. Standard errors clustered at the worker-wave level are in parentheses. The first-stage F-statistic reported in Panel B is the clustered Wald F-statistic for the excluded instrument from the corresponding first stage. *** p<0.01, ** p<0.05, * p<0.1.

variation in the instrument: because tasks vary little within 3-digit occupations, conditioning on them leaves the OAI with insufficient residual variation and yields a weak first stage. We therefore report specifications without occupation controls and with broad (2-digit) occupation controls in the main text, and relegate the 3-digit results to Appendix B for transparency.

Table 6 reports the results. The 2SLS coefficients range from 0.545 - 0.702 (first-stage F scores ranging from 92.97 - 436.37), indicating that a 10 percentage-point increase in

Table 7: Spillovers from High-Adoption Tasks to Non-Work Adoption

	(1) Worker level
Panel A. OLS: Dependent variable = Uses genAI at home	
Uses genAI at work	0.371***
	(0.013)
Observations	11,609
Panel B. 2SLS: Work adoption instrumented by OAI	
Uses genAI at work	0.504***
	(0.129)
First-stage F-stat.	55.38
Observations	11,609
Panel C. First stage: Dependent variable = Uses genAI at work	
Portfolio-average OAI	0.901***
	(0.121)
Observations	11,609

Notes: The dependent variable is an indicator for whether the respondent uses genAI for non-work purposes. Regression is estimated at the worker level. Its work-task AI use variable is an indicator for whether the worker uses genAI, instrumented by the mean leave-one-out adoption rate across the worker's full IWA portfolio. The leave-one-out adoption rates exclude the respondent's own adoption decision and are pooled across the August 2025, November 2025, February 2026, and May 2026 survey waves. Panel A reports OLS estimates, Panel B reports 2SLS estimates, and Panel C reports the corresponding first stage. IWA-level use equals one if the worker uses genAI for any underlying DWA within the IWA. The sample includes workers reporting at least two unique IWAs and only IWAs observed at least 20 times in the pooled data. Controls are included for the number of IWAs reported, a quadratic in age, education, race, sex, survey-wave fixed effects, and broad occupation fixed effects. Regression is weighted using RPS survey weights and includes an unreported constant. Standard errors are heteroskedasticity-robust. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

adoption for a worker's other tasks is associated with a 5.4 - 7.0 percentage point increase in the worker's adoption for the focal-task. The OLS estimates range from 0.621 - 0.628, similar to the 2SLS, suggesting that sorting by workers' underlying propensity to adopt genAI is modest.

Work-home spillovers Our third test asks whether genAI use for work increases adoption outside of work. Through the lens of our framework, for a given worker i, E_i for non-work adoption is proxied by work adoption. Because work adoption may itself reflect the worker's adoption propensity $\bar{\xi}_i$, we instrument for it using the OAI. Here the key endogenous variable

is an indicator for whether the worker reports using genAI at work, instrumented by the average leave-one-out adoption rate of the eligible IWAs in the worker's task portfolio. The regression investigates whether being induced to adopt genAI at work spills over into non-work use. As in the previous exercise, the identifying assumption is that, within a given occupation, the composition of a worker's task portfolio is uncorrelated with the worker's adoption propensity, $\bar{\xi}_i$.

Table 7 shows a strong first stage and a positive relationship in both OLS and 2SLS: workers who use genAI at work are more likely to use genAI for non-work purposes, and the relationship remains positive when work adoption is instrumented with the OAI. Under the identifying assumption, this pattern is consistent with cross-domain learning. Using genAI at work may lead workers to incur the fixed costs of learning and adoption, after which they can apply that knowledge more broadly in non-work activities.

Summary We have shown that some workers are systematically more likely to adopt genAI than others, regardless of the tasks they perform. One potential source of individual differences in adoption is learning from genAI experience: if workers learn how to effectively use genAI in one domain, this may reduce the cost of adopting genAI in other domains. We conducted three empirical tests of this learning channel across different classes of domains: across time, across tasks, and across work versus non-work activities. The results of all three tests were consistent with an important role for learning how to use genAI from experience.

6 Potential Channels of Future GenAI Productivity Gains

The preceding sections document the work that genAI currently performs and the factors that shape its adoption. These findings describe the technology at a particular stage of its diffusion and can inform analyses of the productivity impact of genAI to date. However, these findings can also shed light on how that impact will evolve over time. To distinguish the specific channels through which future productivity gains may emerge, we extend the static framework from Section 3.1 to a dynamic setting.

Equation 3 described the aggregate productivity gain from genAI at a single point in time. A dynamic version of this equation is:

$$\frac{Y(\tau) - Y_0}{Y_0} = \sum_t \omega_t(\tau) \cdot a_t(\tau) \cdot g_t(\tau). \tag{9}$$

Here, genAI productivity, adoption rates, and task shares are a function of time, τ. Within this framework, future gains in productivity from genAI can come from four channels.

Learning-driven diffusion We have shown that individual-specific factors appear to be a primary barrier to genAI adoption, $a_t(\tau)$. However, we also provide evidence that workers are more likely to adopt genAI once they have learned how to effectively use it. This suggests that, over time, workers will gradually adopt genAI for a greater share of their work as they gain experience. The potential for expanded genAI adoption via this channel is substantial given that the vast majority of tasks currently exhibit adoption rates below 50%. Employers may be able to speed up this process through training, coaching, or encouragement (Bick et al., 2026a). In particular, tasks that have high potential for productivity gains, g_t, but which tend to be performed by workers with low propensity to adopt, ξ_i, stand to benefit most from targeted employer support. Instances involving clerical and administrative work, which exhibit low adoption rates relative to predicted genAI potential, are natural candidates.

Lowering task-level adoption barriers Several tasks with high genAI potential but low adoption also involve work that faces compliance requirements, workflow frictions, or confidentiality concerns. Greater adoption in these tasks could be facilitated by reductions in task-specific barriers, μ_t, such as regulatory changes (e.g., HIPAA-compatible deployments), technological solutions (privacy-preserving inference, auditable outputs), or organizational adaptation of workflows.

GenAI improvements Improvements in genAI capabilities can raise productivity through both intensive and extensive margins. First, an increase in $g_t(\tau)$ raises productivity among workers who already use genAI. This channel remains relevant even for tasks with adoption rates approaching 100%, for which little scope remains for gains from further adoption. Second, for tasks with lower adoption rates, an increase in $g_t(\tau)$ may also induce greater adoption; see Section 4.1.1.

Task reallocation Economy-wide task shares $\omega_t(\tau)$ evolve over time in response to technological changes (Autor et al., 2003). If genAI complements human labor in a particular task, we would expect its share of labor to increase as genAI becomes more productive. Alternatively, if genAI substitutes for human labor in a particular task, we would expect its labor share to decrease. Our simple framework is not designed to address this issue, but these compositional shifts represent an important margin for future work.

7 Comparing GenAI Adoption and Chat Data

The RPS measures genAI adoption via workers' self-reports. An alternative approach starts from genAI platform chat logs and classifies chats into specific tasks. In this section, we first describe how chat-based measures map into the conceptual framework introduced in Section 3.1. We show that, on their own, chat-based measures are insufficient to identify aggregate or occupation-level productivity gains from genAI. We then compare our survey-based measures to the chat-based measures in Handa et al. (2025), Chatterji et al. (2025), and Tomlinson et al. (2025).[7] While we find some agreement among the various measures, we also highlight several conceptual and practical challenges that limit their comparability.

7.1 Mapping Chat Data into the Conceptual Framework

Chat data start from a random sample of a platform's genAI chats, $C \in \mathbb{N}$. Each chat is assigned to a single task, resulting in a partition of chats across tasks C_t, where $\sum_t C_t = C$. Task t's share of chats is then $c_t = C_t/C$, with $\sum_t c_t = 1$. We can map these chat-based objects into the economic framework in Section 3.1 by writing the number of task t chats, C_t, as the product of the task's share of work in the economy, ω_t, the adoption rate for the task, a_t, and the number of chats needed to assist one instance of task t, n_t:

$$C_t = \omega_t a_t n_t \tag{10}$$

We can then write task t's share in chat data as

$$c_t = \frac{\omega_t a_t n_t}{C} \tag{11}$$

7.1.1 Conceptual Challenges in Interpreting Chat-Based Data

How do chat shares relate to adoption rates across tasks? Equation (11) implies that the ratio of chat shares for tasks s and t is

$$\frac{c_t}{c_s} = \left(\frac{n_t}{n_s}\right)\left(\frac{\omega_t}{\omega_s}\right)\left(\frac{a_t}{a_s}\right). \tag{12}$$

[7] For OpenAI, we use the February 2026 public OpenAI Signals release; for Anthropic, the March 27, 2025 Anthropic Economic Index release; and for Microsoft,the data from the replication package by Tomlinson et al. (2025). No futher updates have been released. We omit Iscenko et al. (2026) (Google Gemini) from our analysis because task-level data are not publicly available.

Relative chat shares do not generally reflect relative adoption rates unless $(n_t/n_s)\,(\omega_t/\omega_s) = 1$. This expression contains two sets of terms, related to chat intensity, n, and task prevalence, ω.

First, relative chat shares conflate adoption rates with chat intensity when $n_t \neq n_s$. In practice, chat intensity may vary across tasks if genAI assistance in some tasks require many iterative chats, while assistance in other tasks only a single prompt. This latter aspect is likely more relevant for the OpenAI analysis, which defines a chat as a single message.

Second, relative chat shares conflate adoption rates with task prevalence when $\omega_t \neq \omega_s$. Because task shares vary widely (see Figure 5b), recovering relative adoption rates from chat shares requires an external measure of ω_t. One potential approach is to proxy for ω_t by combining ONET task-importance scores with occupational employment shares, which could allow future research using chat data to account more directly for differences in task prevalence. The RPS offers a more direct alternative, containing information on both a_t and ω_t, which allows us to construct a measure of c_t from RPS data that can be compared directly to chat shares.

Do chat shares allow us to estimate the aggregate productivity gain from genAI? Restating Equation (3), the percent change in aggregate productivity from genAI is

$$\frac{\Delta Y}{Y} = \sum_t \omega_t a_t g_t \,.$$

If we assume that chat intensity is constant across tasks, $n_t = n\,\forall t$, the chat share c_t has a natural interpretation in our framework: it is the probability that a randomly drawn genAI-assisted task applies to task t. Applying Bayes' rule,

$$c_t \;=\; \Pr(\text{task} = t \mid \text{genAI used}) \;=\; \frac{\Pr(\text{genAI used} \mid \text{task} = t)\,\Pr(\text{task} = t)}{\Pr(\text{genAI used})} \;=\; \frac{a_t\,\omega_t}{a} \quad (13)$$

where $a \equiv \sum_t \omega_t a_t$ is the economy-wide adoption rate. Substituting Equation (13) after re-arranging terms allows us to rewrite Equation (3) as

$$\frac{\Delta Y}{Y} = a \sum_t c_t g_t \quad (14)$$

The equation states that chat shares c_t can be used to estimate the aggregate productivity gain from genAI if they are combined with estimates of task-level productivity gains g_t and the aggregate adoption rate a. Intuitively, a appears because aggregate gains depend not

only on how genAI use is distributed across tasks, but also on the share of work that genAI assists.[8]

7.1.2 Practical Challenges in Measuring and Comparing Chat-Based Data

In addition to the conceptual challenges discussed above, comparisons of genAI use with chat data also involve several practical challenges.

Definition of a "Chat" Chatterji et al. (2025) define a chat as a single user prompt in OpenAI data and classify chats at the IWA level, using the 10 preceding messages as context. Alternatively, Handa et al. (2025) define a chat as an entire user conversation in Anthropic data and classify chats at the ONET task level, from which we construct an IWA-analogue (see Appendix C for details). Tomlinson et al. (2025) also use the entire Microsoft Copilot conversation as the unit of observation. A Copilot conversation may be assigned to multiple IWAs, in which case its weight is divided equally across the matched IWAs when constructing chat shares. These different methodologies may affect cross-platform comparisons. For example, the larger scope of a chat in Anthropic and Microsoft data may provide additional context relevant for the chat-task assignment procedure.

Non-Representative User Base The users of a given genAI platform may not be representative of the broader economy. For example, a platform's users may differ in their occupational mix, $\tilde{\lambda}_o$, their task mix within occupations, $\tilde{\omega}_{o,t}$, or their adoption rates, $\tilde{a}_t$. In this case, differences in task composition can generate discrepancies between c_t and the economy-wide objects in the framework.

Non-representative user bases could arise for several reasons. One example is that some genAI platforms could have a comparative advantage in particular tasks, and workers who perform these tasks may favor those platforms over others. Another practical example is that both Chatterji et al. (2025) and Handa et al. (2025) exclude users from employer accounts, who may differ systematically from users with personal accounts, see also Yin and Ogut (2026).[9]

Additionally, RPS estimates only capture each worker's top ten tasks, while chat data

[8] An interesting approach is Tamkin and McCrory (2025), which attempts to estimate g_t using the content of chat data. This allows them to estimate the aggregate gain in productivity under the assumption of universal adoption among the subset of genAI-amenable tasks.

[9] More recently, OpenAI and Anthropic released data on enterprise usage.

will capture the full scope of adoption across all tasks. Discrepancies between task shares between chat data and the RPS may also partly reflect differences in coverage rather than true differences in task-level adoption.

Task Classification The procedures above classify each chat to an ONET task, which faces two interrelated challenges. First, chats may omit context necessary to identify which ONET task the chat pertains to. Second, ONET tasks reflect a combination of purposes and basic activities. For example, one of the top ten DWAs for Economics Professors (SOC 25-1063) is "Research topics in area of expertise." This task entails a wide range of basic activities, such as brainstorming, constructing and analyzing economic models, obtaining data, cleaning and analyzing data, coordinating with coauthors, etc. However, each of these basic activities also closely resembles other ONET tasks from other occupations. For example, if an economics professor asks genAI for help analyzing data for a research project, according to ONET this chat should be classified as the DWA "Research topics in area of expertise." However, the classification procedure might classify this task as DWA "Analyze data to identify trends, relationships, or patterns", which is not a DWA for economics professors.

These challenges suggest that chat classification procedures may be predisposed to assigning chats to ONET tasks that reflect basic activities. The next section presents evidence consistent with this hypothesis.

7.2 Comparing Empirical Estimates

Figure 9 compares three chat-based genAI task shares from Chatterji et al. (2025) (OpenAI) Handa et al. (2025) (Anthropic), and Tomlinson et al. (2025) (Microsoft).[10] We also compare genAI task shares from the RPS, which is the share of task t among all AI-assisted worker–task pairs:

$$c_t = \frac{a_t}{a}\omega_t \tag{15}$$

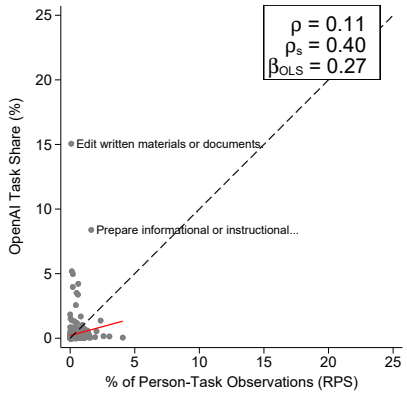
Each of the four measures are vectors of shares summing to one across 332 IWAs.

We highlight two key takeaways. First, the four measures display notable differences in

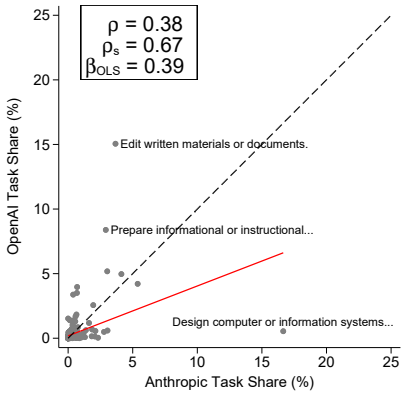
[10] Tomlinson et al. (2025) provide two scores, the "User Goal" and "AI Action" measures. For comparability with the other measures focusing on tasks performed by the worker using genAI, we focus on the "User Goal" measure. Iscenko et al. (2026) also map interactions with Google's Gemini products to ONET tasks, but their task-level data are not publicly available.

Figure 9: Comparing GenAI Task Shares in RPS, OpenAI, Anthropic, and Microsoft Data

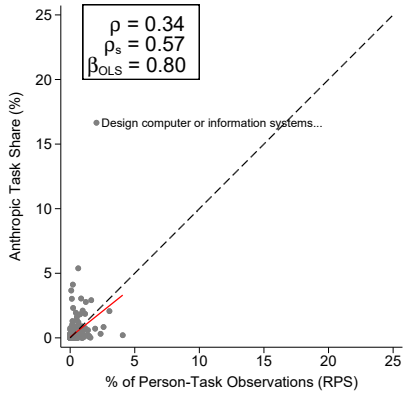
(a) RPS vs. OpenAI Shares



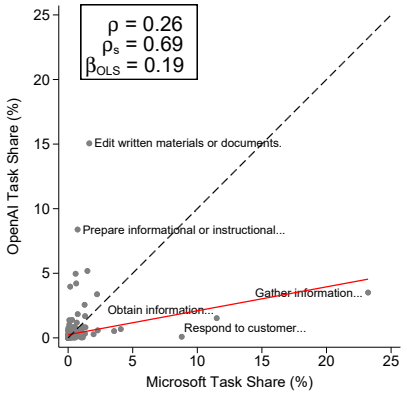
(b) OpenAI vs. Anthropic Shares



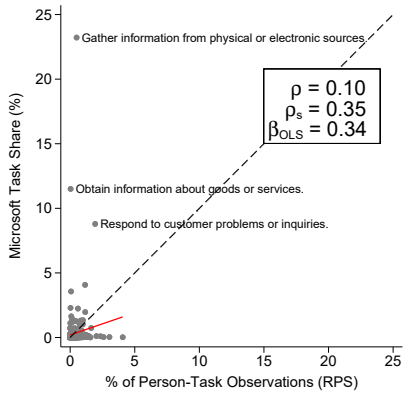
(c) RPS vs. Anthropic Shares



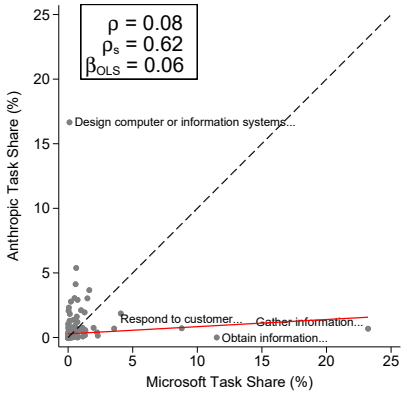
(d) OpenAI vs. Microsoft Shares



(e) RPS vs. Microsoft Shares

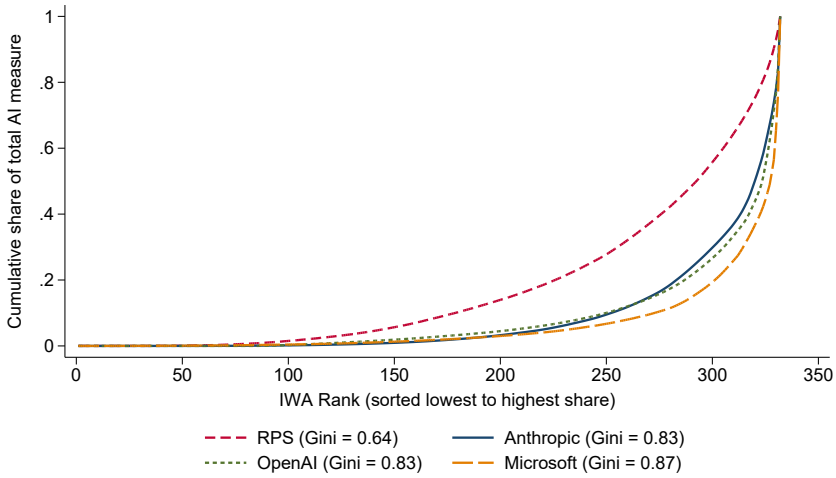


(f) Anthropic vs. Microsoft Shares



Notes: Data for the RPS percent of person-task observations are pooled across the August 2025, November 2025, February 2026, and May 2026 waves of the RPS. Each point is one of 332 IWAs. Dashed gray line is the 45-degree line. Red line is the OLS fitted line. Text boxes report Pearson correlation (ρ), Spearman rank correlation (ρ_s), and the slope of the OLS regression (β_{OLS}). Microsoft measure is based on users' stated goals; results using Microsoft's AI-action classification are similar. Labels are displayed for IWAs which have at least 7% share in one of the datasets used in the graph. For readability of labels, long labels are cut off with "...".

Figure 10: Cumulative Distribution of IWA Task Shares



Notes: Each curve plots the cumulative share of total AI usage accounted for by IWAs, ranked from smallest to largest share. Steeper curves indicate more concentrated usage. The RPS measure (c_t^{RPS}) distributes usage more evenly across tasks than any of the three chat-log measures.

task shares. The correlation between the RPS and chat data sources are 0.11 (OpenAI), 0.34 (Anthropic), and 0.10 (Microsoft). The correlations among the different chat data sources are 0.38 (OpenAI-Anthropic), 0.26 (OpenAI-Microsoft), and 0.08 (Anthropic-Microsoft).

Second, task shares are less concentrated in the RPS. The largest task accounts for 15.1% of use in the OpenAI data, 16.7% in the Anthropic data, and 23.2% in the Microsoft data, compared with 4.1% in the RPS.

Figure 10 confirms this greater concentration in the chat data. For example, the 10 IWAs with the largest task shares (out of 332) together account for 53.0% of use in the OpenAI data, 46.0% in the Anthropic data, 60.8% in the Microsoft data, and only 22.2% in the RPS data.

To investigate the sources of these disparities, Table 8 decomposes the squared differences between the RPS and the three chat-based measures. A single IWA accounts for 50.3% of the total squared disagreement with OpenAI, 60.9% with Anthropic, and 63.5% with Microsoft. The ten largest task-level discrepancies account for 86.6%, 86.7%, and 93.2% of total squared disagreement, respectively. Thus, the disagreement between the RPS and each platform-based measure is driven disproportionately by a small number of tasks with very high shares in the chat-based data.

The tasks driving this disagreement are informative. Table 9 lists the tasks with the largest genAI share in each of the three datasets and in the RPS. The top tasks in the

Table 8: GenAI Task Share Disagreements Are Concentrated in a Small Number of Tasks

(a) All Tasks

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	Microsoft vs. RPS
1	50.3%	60.9%	63.5%
5	74.9%	79.6%	89.0%
10	86.6%	86.7%	93.2%
$\sum_t d_t^2$	0.0446	0.0351	0.0813
# IWAs	332	332	332

(b) Tasks Overrepresented in Chat Data

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	Microsoft vs. RPS
1	59.0%	71.1%	70.0%
5	87.4%	90.8%	97.1%
10	97.3%	95.9%	98.9%
$\sum_t d_t^2$	0.0380	0.0301	0.0737
# IWAs	96	100	109

(c) Tasks Underrepresented in Chat Data

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	Microsoft vs. RPS
1	24.5%	29.7%	21.7%
5	54.9%	52.6%	54.0%
10	67.0%	65.4%	66.7%
$\sum_t d_t^2$	0.0066	0.0050	0.0076
# IWAs	219	219	222

Notes: Panel (a) decomposes total squared disagreement $\sum_t d_t^2$ across all 332 IWAs. Panels (b) and (c) decompose squared disagreement separately among IWAs where chat logs assign a higher share than the RPS ($d_t > 0$, "overrepresented") and a lower share ($d_t < 0$, "underrepresented"). Rows show the cumulative share of the relevant $\sum d_t^2$ explained by the top k IWAs ranked by d_t^2. Bottom rows report the total $\sum d_t^2$ and the number of IWAs in each group.

chat data sources all describe basic, generic actions that apply across many contexts: "Edit written materials or documents," "Design computer or information systems or applications," and "Gather information from physical or electronic sources." These tasks all have fairly high adoption rates in the RPS, but only a modest share of workers are in occupations that contain these tasks in ONET. By contrast, the task with the largest genAI share in the RPS is "Direct organizational operations, activities, or procedures," which is a higher-order, purpose-oriented title that encompasses several more basic actions such as brainstorming, analysis, and writing.

Table 9: Top GenAI Task from Each Data Source

IWA	Share of genAI uses (%)				RPS adoption (%)	CPS prevalence (%)
	RPS	OpenAI	Anthropic	Microsoft		
Edit written materials or documents.	0.1	**15.1**	3.7	1.6	45.1	2.4
Design computer or information systems or applications.	2.0	0.5	**16.7**	0.1	51.9	11.5
Gather information from physical or electronic sources.	0.5	3.5	0.7	**23.2**	35.1	6.6
Direct organizational operations, activities, or procedures.	**4.1**	0.0	0.2	0.0	33.5	43.0

Notes: Rows are the union of the highest-share IWA in the RPS, OpenAI, Anthropic, and Microsoft data; bold values identify the maximum within each source. The first four columns report the percentage of genAI uses assigned to each IWA, with shares normalized to sum to 100 across all 332 IWAs. The RPS shares pool the August 2025, November 2025, February 2026, and May 2026 survey waves and use RPS survey weights. The Microsoft measure uses the user-goal classification. RPS adoption is the weighted percentage of respondents performing the IWA who report using genAI for it. CPS prevalence is the employment-weighted percentage of CPS workers whose occupation contains the IWA in ONET.

As described in Section 7.1.2, a natural interpretation of these results is that chat classifiers may be predisposed to assigning chats to task titles describing basic, generic actions. For example, over 15% of chats in OpenAI data are classified as "Edit written materials or documents," even though only 2.4% of workers are in occupations that include this task in ONET. While the share of workers who perform editing is clearly much larger than 2.4%, in most occupations ONET implicitly embeds editing within higher-order, purpose-oriented tasks, such as preparing reports, drafting legal documents, or conducting research. But because chat classifiers do not account for a user's occupation, tasks with basic, generic titles may act as catch-all categories for classifiers analyzing chats from a variety of occupational contexts. This explains why editing has a modest task share in the RPS and a much higher share in chat data.[11]

Table 8 is consistent with this interpretation. Panel (b) shows that, among tasks overrepresented in chat data, a few tasks drive nearly all the total differences in chat shares between the RPS and chat shares. Alternatively, Panel (c) shows that underrepresented tasks are more diffuse, with disagreement spread across many tasks. This asymmetry is consistent with a small set of basic, generic task titles absorbing chats that ONET would map to a broader set of higher-order, purpose-oriented tasks. (By contrast, disagreements between the chat data sources find similarly concentrated differences for both under- and

[11]Recent research based on classifying genAI chats have taken these considerations into account. For example, Chin and Richmond (2026) label broadly shared tasks such as editing written materials as "generic" rather than occupation-specific.

Table 10: GenAI Task Share Correlations For Various Aggregation Levels

Panel A: RPS versus Chat Measures				
	N	OpenAI vs. RPS	Anthropic vs. RPS	Microsoft vs. RPS
IWA	332	0.11	0.34	0.10
WA	37	0.54	0.54	0.21
BWA	9	0.58	0.56	0.20
Panel B: Comparisons Among Chat Measures				
	N	OpenAI vs. Anthropic	OpenAI vs. Microsoft	Anthropic vs. Microsoft
IWA	332	0.38	0.26	0.08
WA	37	0.72	0.37	0.17
BWA	9	0.94	0.26	0.04

Notes: Pearson correlations of task shares across data sources at three levels of aggregation: Intermediate Work Activities (IWA, 332 categories), Work Activities (WA, 37), and Broad Work Activities (BWA, 9). Panel A compares the RPS with each chat-based measure; Panel B reports pairwise correlations among the three chat-based measures. N denotes the number of task categories at each level.

over-represented tasks.)

Finally, a related question is whether these disagreements also reflect fine-grained classification noise among similar tasks that washes out at higher levels of aggregation, or more fundamental differences in what the data sources capture. Table 10 shows that the correlation in task shares across datasets tends to increase when tasks are aggregated from 332 IWAs to 37 WAs or 9 BWAs. The exception is the Microsoft dataset, which shows fairly consistent disagreement with other datasets even at higher levels of aggregation. Aggregation therefore resolves some fine-grained classification disagreement, but substantial differences remain even across nine broad task categories.

Summary Taken together, this section documents systematic differences across chat-based measures of genAI use and between chat-based measures and the RPS. A central source of disagreement is how chats are assigned to tasks. Chat-based measures assign tasks based on chats themselves, and tend to assign usage to a small number of tasks with generic, activity-based titles. In contrast, the RPS maps genAI use to ONET tasks tied to occupations, including higher-order, purpose-oriented tasks that are underrepresented in chat data. This difference leads to highly concentrated chat shares and more diffuse survey-based shares, with a small number of generic tasks accounting for most of the disparities. Aggregation to broader tasks reduces but does not eliminate disagreement between the RPS and chat-based data, suggesting that fine-grained classification differences explain only part of the

divergence. Other differences across datasets, such as user populations, may also contribute, though we are not able to quantify their relative importance.

Over-classification of generic, activity-based tasks (relative to the ONET framework) highlight two key considerations when interpreting this data. First, survey-based genAI task shares, which start from a respondent's occupation and ask about specific ONET tasks, are not cleanly comparable to chat-based task shares, which may lack the occupation- and purpose-based context needed to classify a chat in a manner consistent with ONET. This suggests that chat classification procedures may benefit from alternative task frameworks, potentially outside of ONET, that place greater emphasis on the underlying basic activities performed rather than purpose-based frameworks. Second, occupation shares constructed from chat-task classifications are likely to be inaccurate. For example, Handa et al. (2025) construct occupation genAI shares by mapping chats to tasks and then mapping tasks to occupation using ONET. (These shares have then been used as proxies for genAI exposure in other research, e.g. Brynjolfsson et al., 2026; Edlich and Slok, 2026; Gimbel et al., 2025; Massenkoff and McCrory, 2026). But if generic tasks capture chats from workers whose occupations do not include that task in ONET, this procedure will misattribute genAI use across occupations.

8 Conclusion

This paper presents novel measures of what work genAI does, rather than predicting what work it might do. GenAI adoption varies substantially across occupations and tasks. We organize these patterns into occupation- and task-level adoption indexes that can be used to study genAI's labor market impact. A striking feature of current adoption is that it is widespread but shallow: genAI reaches over 80% of occupations and over 40% of tasks, but within most of these fewer than half of workers adopt. This implies substantial adoption variation among workers performing similar tasks. While exposure scores and chat shares can speak to adoption differences across tasks, they are largely silent about worker-level adoption differences within tasks: exposure scores are defined at the task level, not the worker level, while chat shares have no information about non-adopters. But our results suggest that understanding why some workers adopt and others do not, even when they perform similar work, may matter as much for genAI's labor market impact as understanding which tasks genAI can do.

References

Autor, D. H., F. Levy, and R. J. Murnane (2003). "The skill content of recent technological change: An empirical exploration". In: *The Quarterly journal of economics* 118.4, pp. 1279–1333.

Bick, A. and A. Blandin (2023). "Employer reallocation during the COVID-19 pandemic: Validation and application of a do-it-yourself CPS". In: *Review of Economic Dynamics* 49, pp. 58–76.

Bick, A., A. Blandin, A. Caplan, and T. Caplan (2025a). "Heterogeneity in Work From Home: Evidence from Six U.S. Datasets". In: *Federal Reserve Bank of St. Louis Review* 107.14, pp. 1–23.

Bick, A., A. Blandin, A. Caplan, and T. Caplan (2025b). "Measuring Trends in Work From Home: Evidence from Six U.S. Datasets". In: *Federal Reserve Bank of St. Louis Review* 107.15, pp. 1–23.

Bick, A., A. Blandin, D. J. Deming, N. Fuchs-Schündeln, and J. Jessen (2026a). *Mind the Gap: AI Adoption in Europe and the US*. Tech. rep. National Bureau of Economic Research.

Bick, A., A. Blandin, and D. J. Deming (2026b). "The Rapid Adoption of Generative AI". In: *Management Science*. Published online January 20, 2026.

Brynjolfsson, E., B. Chandar, and R. Chen (2026). *Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence*. Working Paper.

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). "What can machines learn and what does it mean for occupations and the economy?" In: *AEA papers and proceedings*. Vol. 108. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 43–47.

Chatterji, A., T. Cunningham, D. J. Deming, Z. Hitzig, C. Ong, C. Y. Shan, and K. Wadman (2025). *How people use chatgpt*. Tech. rep. National Bureau of Economic Research.

Chin, C. and A. M. Richmond (July 2026). *Work at the Frontier: How AI Is Expanding What People Do at Work*. Tech. rep. OpenAI.

Deming, W. E. and F. F. Stephan (1940). "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known". In: *The Annals of Mathematical Statistics* 11.4, pp. 427–444.

Edlich, S. and T. Slok (2026). *The Impact of AI on the U.S. Labor Market: Early Evidence from Observed Adoption*. Apollo Global Management White Paper, July 2026.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2024). "GPTs are GPTs: Labor market impact potential of LLMs". In: *Science* 384.6702, pp. 1306–1308.

Felten, E., M. Raj, and R. Seamans (2021). "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses". In: *Strategic Management Journal* 42.12, pp. 2195–2217.

Fletcher, R. and R. Nielsen (2024). *What does the public in six countries think of generative AI in news?* Reuters Institute for the Study of Journalism.

Gimbel, M., M. Kinder, J. Kendall, and M. Lee (2025). *Evaluating the Impact of AI on the Labor Market: Current State of Affairs*. Tech. rep. October 2025; updated through 2026 with monthly CPS and Anthropic usage releases. The Budget Lab at Yale.

Handa, K., A. Tamkin, M. McCain, S. Huang, E. Durmus, S. Heck, J. Mueller, J. Hong, S. Ritchie, T. Belonax, K. K. Troy, D. Amodei, J. Kaplan, J. Clark, and D. Ganguli (2025). *Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations*. arXiv: 2503.04761 [cs.CY].

Iscenko, Z., S. Strand, Y. Chen, G. Aimard, M. Codreanu, V. Sampathkumar, A. Imas, J. Jacobs, E. Muiruri, J. Mateos-Garcia, J. J. Ng, S. Javed, J. Martin, O. Ajmeri, D. Calin, A. Kim, F. C. Millet, and J. Manyika (2026). *Google's AI Economy ATLAS v1.0: Mapping Gemini Usage in the Economy*. arXiv: 2608.00038 [econ.GN].

Laughlin, L., X. Song, M. Wisniewski, and J. Xu (2024). *From Job Descriptions to Occupations: Using Neural Language Models to Code Job Data*. Working Paper. U.S. Census Bureau, University of Pennsylvania, and Pennsylvania State University.

Lindenlaub, I., R. Oh, M. A. Rodriguez, and L. Veldkamp (2026). *Beyond exposure: predicting AI adoption based on comparative advantage*. Tech. rep. National Bureau of Economic Research.

Massenkoff, M. and P. McCrory (2026). *Labor Market Impacts of AI: A New Measure and Early Evidence*. Anthropic, March 2026.

McClain, C. (2024). *Americans' use of ChatGPT is ticking up, but few trust its election information*. Pew Charitable Trusts.

Miano, A. (Jan. 2025). *Search Costs, Outside Options, and On-the-Job Search*. Working Paper. University of Naples Federico II and CSEF.

Tamkin, A. and P. McCrory (Nov. 5, 2025). *Estimating AI productivity gains from Claude conversations*. URL: https://www.anthropic.com/research/estimating-productivity-gains.

Tomlinson, K., S. Jaffe, W. Wang, S. Counts, and S. Suri (2025). "Working with AI: Measuring the Applicability of Generative AI to Occupations". In: *arXiv preprint arXiv:2507.07935*.

US Census Bureau (2015). "Current Population Survey Interviewing Material". In: *https://www2.census.gov/programs-surveys/cps/methodology/intman/CPS_Manual_April2015.pdf*.

Webb, M. (2020). *The Impact of Artificial Intelligence on the Labor Market*. Working Paper. Stanford University.

Yin, M. and B. Ogut (2026). *Who Uses AI? Platform Selection and the Measurement of Occupational AI Exposure*. arXiv:2605.21743.

Yin, M., H. Vu, and C. Persico (Apr. 2026). *How (un)Stable Are LLM Occupational Exposure Scores? Evidence from Multi-Model Replication*. NBER Working Paper 35110. Cambridge, MA: National Bureau of Economic Research.

ONLINE APPENDIX

A RPS: Measurement and Definitions

A.1 Sample Restrictions

The Qualtrics panel includes about 15 million members and is not a random sample of the U.S. population. However, researchers can instruct Qualtrics to target survey invitations to specific demographic groups. The RPS sample was designed to be nationally representative of the U.S. across gender, age, education, household income, and other demographic characteristics. Individuals who take the survey too fast, i.e. in less than 50% of the median time in a soft launch, or who do not state that they will provide their best answers, are automatically dropped from the sample.[1] Finally, once fielding is completed Qualtrics staff goes over all responses and filters out any that look suspicious. We also screened out responses based on various open text questions who appear to be bots.

Table A.1 compares the sample composition between the CPS and RPS along the demographics targeted in the sampling procedure for our main surveys (columns 1 and 2). The bottom panel of Table A.1 compares employment status in the CPS and RPS, statistics that have not been targeted in the sampling procedure. Individuals classified as employed at work and unemployed are overrepresented in the RPS.

A.2 Weighting

As described in the body of the paper, we asked Qualtrics to administer the survey to a sample of respondents who match the U.S. population along a few broad demographic characteristics: gender, five age bins (18-24, 25-34, 35-44, 45-54, 55-64), race and ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, other), education (high school or less, some college or associate degree, bachelor's degree or more), marital status (married or not), number of children in the household (0, 1, 2, 3 or more), three income bins for household income over the last 12 months (<\$50k, \$50k-100k, >\$100k), and four Census regions.

[1] The exact phrasing of the screener question is: "We care about the quality of our survey data and hope to receive the most accurate measure of your opinions, so it is important to us that you thoughtfully provide your best answer to each question in the survey. Do you commit to providing your thoughtful and honest answers to the questions in this survey?" with the following three answer options (1) I will provide my best answers, (2) I will not provide my best answers, and (3) I can't promise either way. According to Qualtrics staff, this question provides the best results in terms of screening out respondents.

Table A.1: Sample Composition in the Pooled CPS and RPS

	Everyone		*Employed*	
	CPS	RPS	CPS	RPS
	(1)	(2)	(3)	(4)
Gender: Women	50.6	51.5	47.5	47.4
Age				
18–24	15.1	12.9	11.9	13.4
25–34	22.5	20.4	24.2	22.0
35–44	22.3	26.1	24.7	27.8
45–54	19.8	23.4	21.5	22.8
55–64	20.2	17.2	17.7	13.9
Race/Ethnicity				
Non-Hispanic White	54.7	55.9	55.9	56.2
Non-Hispanic Black	12.4	16.2	11.6	16.0
Hispanic	21.4	19.9	21.2	19.8
Other	11.5	8.1	11.3	7.9
Education				
High school or less	36.3	31.1	32.1	26.8
Some college/associate's degree	25.9	28.1	25.3	27.4
Bachelor's or graduate degree	37.8	40.8	42.6	45.7
Marital Status: Married	49.8	49.6	52.7	52.7
Number of children under age 18				
0	61.4	57.2	60.4	54.0
1	17.5	18.9	17.8	20.3
2	13.6	15.2	14.5	16.7
3+	7.5	8.8	7.3	9.0
Household Income in Last 12 Months				
\$0–50,000	23.8	32.0	18.0	24.6
\$50,000–100,000	29.3	26.9	29.9	28.7
\$100,000–150,000	18.7	17.5	20.4	19.9
\$150,000+	28.2	23.6	31.7	26.8
Region				
Northeast	17.1	18.8	17.2	19.8
Midwest	20.2	17.4	20.9	16.9
South	38.8	39.9	38.0	39.3
West	24.0	23.9	23.9	24.0
Employment Status				
Employed, at work last week	72.1	74.2		
Employed, absent from work last week	2.2	2.8		
Unemployed	3.4	7.5		
Not in the labor force	22.4	15.5		
Observations	212304	19030	158472	14655

Notes: Column 1 reports the sample composition in the pooled August 2025, November 2025, February 2026, and May 2026 Current Population Survey (CPS) for the variables targeted by Qualtrics in the sampling procedure. The employment status was the only variable not targeted. Column 2 reports the sample composition in the pooled August 2025, November 2025, February 2026, and May 2026 Real-Time Population Survey (RPS). The sample in both data sets is restricted to the civilian population ages 18-64. Columns 3 and 4 report the same outcomes for the employed (at work and absent from work last week).

Table A.1 compare the sample composition between the pooled August 2025, November 2025, February 2026, and May 2026 RPS and the CPS along the demographics targeted in the sampling procedure for our main survey. The tables look very similarly for each survey wave separately, including the June pilot, and are available upon request. As discussed in the paper, we also compare at the bottom of the table employment status in the CPS and RPS, as well as the demographic composition for the employed. None of this latter sets of moments were targeted during the sampling of survey respondents.

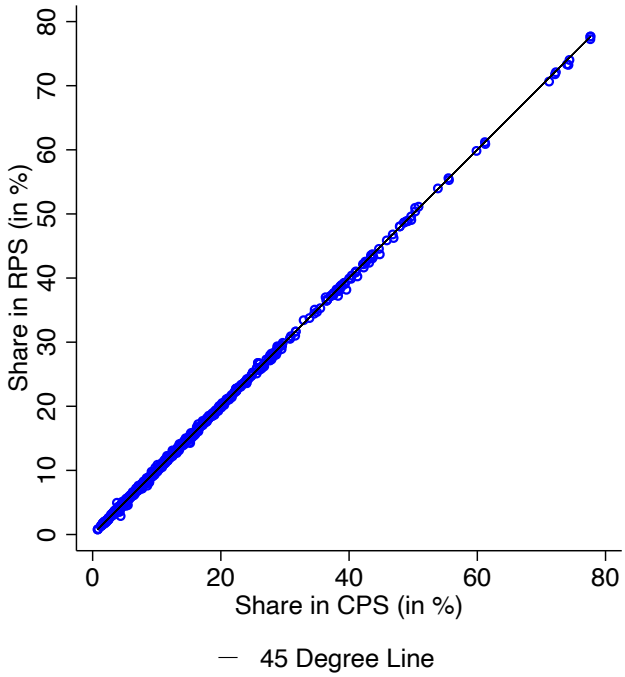
Using the iterative proportional fitting (raking) algorithm of Deming and Stephan (1940), we construct sampling weights to ensure the RPS matches the CPS sample proportions for the same set of demographic characteristics included in the Qualtrics sampling targets for the overall sample, i.e., independent of employment status. However, we use more disaggregated categories for education, marital status, and household income and we interact all categories with gender. In particular, for education, we distinguish between less than high school, high school graduate or equivalent, some college but no degree, associate degree, bachelor's degree, and graduate degree. For marital status, we distinguish between married + spouse present, divorced, never married, and "other." We also condition on relationship status (spouse living in the same household, partner living in the same household, other). For household income over the last 12 months we distinguish between \$0-\$25,000, \$25,000-\$50,000, \$50,000-\$75,000, \$100,000-\$150,000, and more than \$150,000.

In addition, our sampling weights replicate the employed-at-work rates, the employment rates, and the labor force participation rates in each of the subsequent months. We match these key labor market statistics not only in the aggregate but also conditional on demographic characteristics. More specifically, we match the employed-at-work rate, the employment rate, and the labor force participation rate for the current month by the same demographics described in the previous paragraph. We also include 2-digit occupation codes in our weighting scheme.

We also include occupation into our weighting scheme. Among the 22 broad occupations, we merge a) "Community and Social Service Occupations" and "Legal Occupations", b) "Healthcare Support Occupations" and "Protective Service Occupations", and c) "Farming, Fishing, and Forestry Occupations" and "Construction and Extraction Occupations." to ensure a sufficient sample size

Including all the interaction terms, we have a total of 48 statistics (e.g., gender is one, and gender x education is another) which we weight on, with a combined 295 categories (e.g., gender has two categories, and gender x education has 2 x 6 = 12 categories. To visualize

FIGURE A.1: Weighted Sample Composition in the Pooled August 2025, November 2025, February 2026, and May 2026 CPS and RPS



Notes: The figure shows for all statistics used in the weighting scheme, the weighted fraction of individuals with the respective characteristics in the RPS (on the y-axis) and the CPS (on the x-axis).

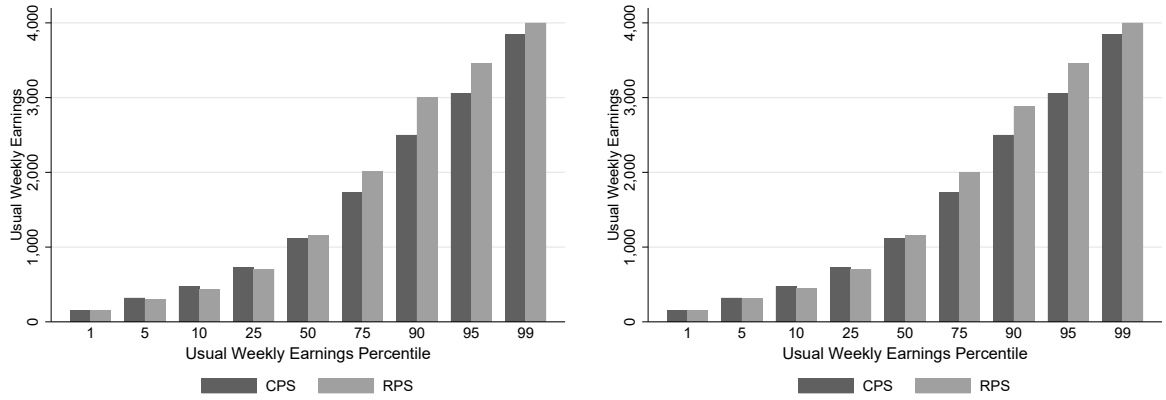
the goodness of fit, we plot in Figure A.1 for all statistics used in the weighting scheme, the weighted fraction of individuals with the respective characteristics in the RPS (on the y-axis) and the CPS (on the x-axis) for August 2025, November 2025, February 2026, and May 2026. The lack of sizeable deviations from the 45-degree line demonstrates how well the weighting procedure works. The results are similar for each survey wave separately, including the June pilot, and are available upon request.

A.3 Validation Checks

To illustrate how the RPS and CPS compare for characteristics that are not part of the sampling scheme and the effect of weighting, panels (a) and (b) of Figure A.2 compare the usual weekly earnings distribution in the RPS and CPS in the pooled August 2025, November 2025, February 2026, and May 2026 waves, with and without weights, respectively. Both the unweighted and weighted distributions are similar, with the RPS featuring somewhat higher earnings in the upper part of the weekly earnings distribution.

FIGURE A.2: Validation Checks – Usual Weekly Earnings

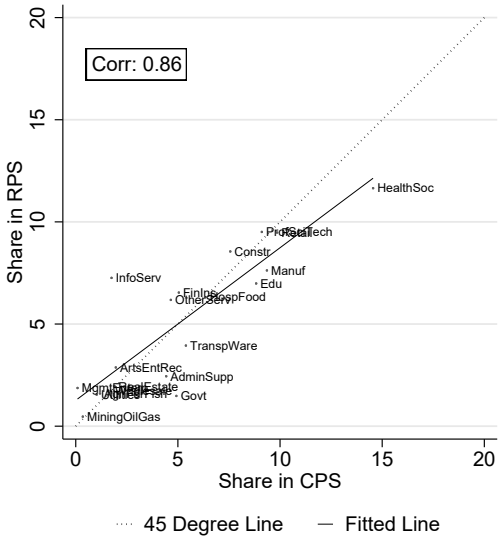
(a) Weekly Earnings Percentiles: Unweighted (b) Weekly Earnings Percentiles: Weighted



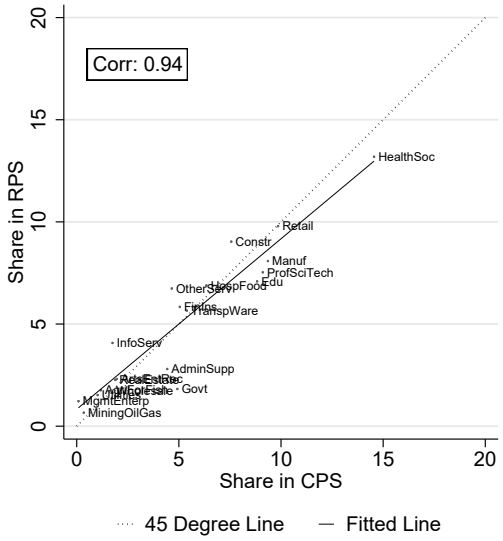
Notes: The figure on the left uses unweighted RPS data, the figure on the right uses weighted RPS data. We use the same sample of RPS respondents in both figures. All figures use weighted CPS data. Data samples for the weekly earnings figures are employees ages 18-64 in the pooled August 2025, November 2025, February 2026, and May 2026 RPS and CPS-ORG with weekly earnings below the CPS topcode of $4300 and an implied hourly wage of at least the federal minimum wage of $7.25. Since the CPS topcode varies across month, we apply the lower topcode in all months.

FIGURE A.3: Validation Checks – Industry and Major Shares

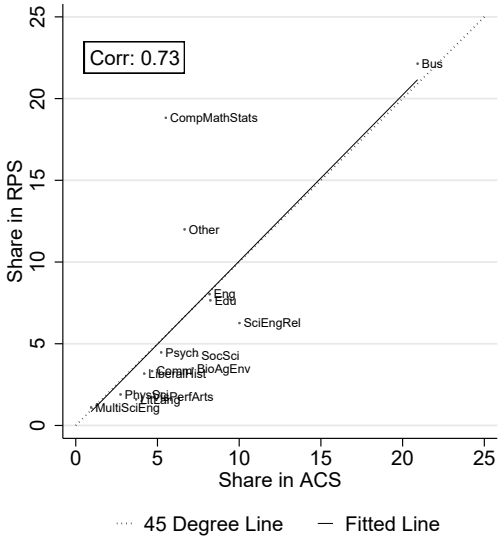
(a) Industry: Unweighted



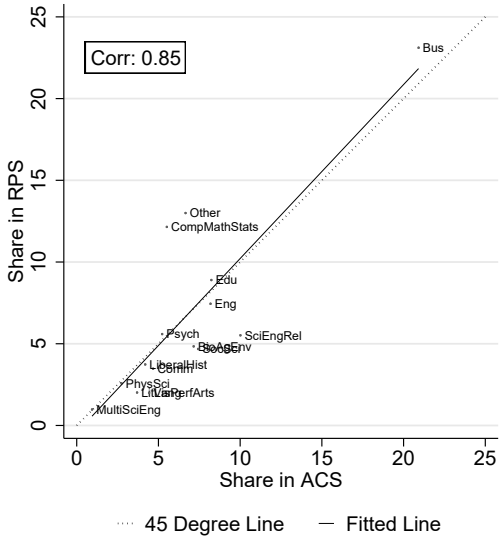
(b) Industry: Weighted



(c) Major Shares: Unweighted



(d) Major Shares: Weighted



Notes: Figures on the left use unweighted RPS data, figures on the right use weighted RPS data. We use the same sample of RPS respondents in both figures. All figures use weighted CPS and ACS data, respectively. Data samples for the industry comparison are employed respondents ages 18-64 in the pooled August 2025, November 2025, February 2026, and May 2026 RPS and CPS with sample sizes of 13920 and 158472, respectively. Data samples for the college major comparison are respondents with a college degree ages 18-64 in the pooled August 2025, November 2025, February 2026, and May 2026 RPS with sample sizes of 7693 and 158472 in the 2022 ACS, respectively.

Figure A.3 compare the industry shares in the RPS and CPS in the pooled August 2025, November 2025, February 2026, and May 2026 waves, with and without weights, as well as college major shares in the RPS and 2022 American Community Survey (ACS). The unweighted correlations of the industry shares is 0.86. The correlation of college majors in the RPS and the ACS is 0.73. The major with the highest share relative to the ACS is Computers, Mathematics, & Statistics, while the major with the lowest share relative to the ACS is Science and Engineering Related Fields. Weighting the data improves the fit between the RPS and the two other data sets.

A.4 Definition of Industries and Occupations

Industries correspond to the 22 major industries in the NAICS: Agriculture, Forestry, Fishing, and Hunting (NAICS=11); Mining, Quarrying, and Oil and Gas Extraction (NAICS=21); Utilities (NAICS=22); Construction (NAICS=23); Manufacturing (NAICS=31-33); Wholesale Trade (NAICS=42); Retail Trade (NAICS=44-45); Transportation and Warehousing (NAICS=48-49); Information (NAICS=51); Finance and Insurance (NAICS=52); Real Estate and Rental and Leasing (NAICS=53); Professional, Scientific, and Technical Services (NAICS=54); Management of Companies and Enterprises (NAICS=55); Administrative and Support and Waste Management Services (NAICS=56); Educational Services (NAICS=61); Health Care and Social Assistance (NAICS=62); Arts, Entertainment, and Recreation (NAICS=71); Accommodation and Food Services (NAICS=72); Other Services (except Public Administration) (NAICS=81); Public Administration (NAICS=99).

The industry groupings used in the main text figures are: Agriculture / Extraction (sectors 11, 21), Construction (23), Manufacturing (31-33), Wholesale / Retail Trade (42, 44-45), Transportation / Utilities (22, 48-49), Info Services (51), Finance / Real Estate (52, 53), Professional / Business Services (54, 55, 56), Education / Health/ Social / Public Services (61, 62, 92), Leisure / Accommodation / Other (71, 72, 81).

Occupations correspond to the 22 major occupations in the SOC: Management Occupations (SOC=11); Business and Financial Operations Occupations (SOC=13); Computer and Mathematical Occupations (SOC=15); Architecture and Engineering Occupations (SOC=17); Life, Physical, and Social Science Occupations (SOC=19); Community and Social Service Occupations (SOC=21); Legal Occupations (SOC=23); Educational Instruction and Library Occupations (SOC=25); Arts, Design, Entertainment, Sports, and Media Occupations (SOC=27); Healthcare Practitioners and Technical Occupations (SOC=29); Healthcare Support Occupations (SOC=31); Protective Service Occupations (SOC=33); Food

Preparation and Serving Related Occupations (SOC=35); Building and Grounds Cleaning and Maintenance Occupations (SOC=37); Personal Care and Service Occupations (SOC=39); Sales and Related Occupations (SOC=41); Office and Administrative Support Occupations (SOC=43); Farming, Fishing, and Forestry Occupations (SOC=45); Construction and Extraction Occupations (SOC=47); Installation, Maintenance, and Repair Occupations (SOC=49); Production Occupations (SOC=51); Transportation and Material Moving Occupations (SOC=53).

A.5 Text of CPS Computer and Internet Use Supplement Questions

In 1984, 1989, 1993, 1997, 2001, and 2003, the leading questions in the CPS Computer and Internet Use supplement (CIU) about computer adoption were:

Employed Respondents: *Do you [directly] use a computer for your job? (No/Yes)*

All Respondents: *Do you [directly] use a computer at home? (No/Yes)*

From 2001-on, the questions omit the word "directly"; we follow this version. While the CIU asks about computer use "at home", we use broader language to account for genAI use on mobile devices outside the home.

A.6 Measurement of O*NET Tasks and Activities

ONET task statements are detailed descriptions of the activities that workers in various occupations perform. ONET comprises 19,530 task statements, which are suggested by actual workers in each occupation who complete surveys indicating which tasks they perform. Workers also rate each task's importance on a 1-5 scale; these ratings are averaged across all surveyed workers in that occupation to create an importance score on a scale from 0 to 100.

To access the importance scores, we download the tasks and importance scores for each 2018 SOC code. We use 2018 codes to stay consistent with the NIOSH Industry and Occupation Computerized Coding System that we use when eliciting respondents' occupations. Some occupations are slightly aggregated. For example, the 2019 SOC lists occupation 11-1011.00 (Chief Executives) and 11-1011.03 (Chief Sustainability Officers) separately; according to the ONET-SOC 2019 to 2018 crosswalk, both map to 2018 SOC code 11-1011 (Chief Executives) and are thus assigned the tasks and importance scores from that code. This procedure yields 15,355 ONET tasks, with importance scores unavailable for 727.

We prefer using O*NET work activities rather than task statements for several reasons. Task statements are long, wordy, and often technical, while detailed work activities are more intuitive. For example:

> ***Task Statement:*** *Typeset and measure dimensions, spacing, and positioning of page elements, such as copy and illustrations, to verify conformance to specifications, using printer's ruler or layout software.*
>
> ***Work Activity:*** *Proofread documents, records, or other files to ensure accuracy.*

> ***Task Statement:*** *Analyze job orders, drawings, blueprints, specifications, printed circuit board pattern films, and design data to calculate dimensions, tool selection, machine speeds, and feed rates.*
>
> ***Work Activity:*** *Study blueprints or other instructions to determine equipment setup requirements.*

Using crosswalks from O*NET, we link 13,803 of the 15,355 task statements with importance scores to detailed work activities; 10,659 have a unique match, and for the remaining 3,144, most link to two activities. When multiple task statements link to a single occupation-specific detailed work activity, we assign the mean importance score.

The 2,040 detailed work activities can be aggregated into 332 intermediate work activities, 37 broader work activities, 9 broad work activities, and 4 very broad work activities. For 25 occupations, we do not have importance scores for any designated tasks or work activities. Based on the respondent's occupation, we ask:

> *Below is a list of work activities commonly associated with your occupation. Please select all the activities that you personally perform as part of your job.*

We present respondents with the top 10 activities by importance, except for 57 occupations with fewer than 10 activities (affecting 48 respondents in August 2025 and 50 in November 2025). GenAI users are then asked:

> *Earlier in the survey you indicated that you personally perform the following activities in your job. For which of these activities do you regularly use Generative AI? Please select all that apply.*

B Cross-Task Spillovers with Detailed Occupation Controls

The main analysis reports specifications without occupation controls and with broad occupation controls. More detailed occupation controls weaken the required assumption about sorting across occupations, but they also absorb substantial variation in the OAI instrument because workers within detailed occupations tend to perform similar task portfolios. Table B.2 additionally reports a specification with detailed occupation controls and shows that this problem of absorption of variation in the OAI instrument leads to a weak first stage. We therefore report this specification for transparency but note that the 2SLS estimate is unreliable given the weak instrument.

Table B.2: Other-Task GenAI Use and Focal-Task Adoption: Alternative Occupation Controls

Occupation controls	(1) None	(2) Broad	(3) Detailed
Panel A. OLS: Dependent variable = Uses genAI for focal task			
Other-task AI use share	0.628***	0.621***	0.590***
	(0.011)	(0.011)	(0.011)
Observations	55091	55091	55091
Panel B. 2SLS: Other-task AI use share instrumented by OAI			
Other-task AI use share	0.702***	0.545***	0.129
	(0.039)	(0.067)	(0.342)
First-stage F-stat.	436.37	92.97	6.73
Observations	55091	55091	55091
Panel C. First stage: Dependent variable = Other-task AI use share			
Other-task adoption index (OAI)	0.818***	0.615***	0.220***
	(0.039)	(0.064)	(0.085)
Observations	55091	55091	55091

Notes: The unit of observation is a worker-IWA pair. The dependent variable in Panels A and B indicates whether the worker uses genAI for the focal IWA. Other-task AI use share is the share of the worker's other reported IWAs for which the worker uses genAI. The Other-Task Adoption Index (OAI) is the leave-one-out mean adoption rate of the worker's other IWAs, excluding the worker's own adoption decisions. To construct IWA-level outcomes, the data are collapsed to unique worker-IWA observations, with genAI use equal to one if the worker uses genAI for any underlying DWA within the IWA. The sample includes only IWAs observed at least 20 times in the pooled August 2025, November 2025, February 2026, and May 2026 survey waves. All regressions control for education, race, sex, a quadratic in age, survey-wave fixed effects, the number of IWAs reported by the worker, and focal-IWA fixed effects. Columns (1), (2), and (3) respectively include no occupation controls, broad occupation controls, and detailed occupation controls. Regressions are weighted using RPS survey weights and include an unreported constant. Standard errors clustered at the worker-wave level are reported in parentheses. The first-stage F-statistic is the clustered Wald F-statistic for the excluded instrument. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C Construction of IWA-level Anthropic Measure

Anthropic's public Economic Index data are released at the O*NET task-statement level rather than at the Intermediate Work Activity (IWA) level. The March 27, 2025 release includes `task_pct_v2.csv`, which reports a task-level prevalence measure (`task_name`, `pct`),

and `onet_task_statements.csv`, which contains O*NET task statements and their associated O*NET-SOC occupation codes. Anthropic's data dictionary defines `task_name` as the normalized task name and `pct` as the percentage prevalence of that task in the dataset. Anthropic also provides a replication notebook, `v2_report_replication.ipynb`, which merges `task_pct_v2.csv` to `onet_task_statements.csv` and then aggregates to occupations.

Because our analysis is conducted at the IWA level, we map Anthropic's task-level measure into the O*NET work-activity hierarchy. We do this in two steps. First, we map Anthropic tasks to Detailed Work Activities (DWAs) using O*NET's `Tasks to DWAs` file. O*NET states that each DWA is mapped to multiple task statements and that each referenced task statement is mapped to one or more DWAs, so this bridge is many-to-many. Second, we map DWAs to Intermediate Work Activities (IWAs) using O*NET's `DWA Reference` file. O*NET states that each DWA is linked to exactly one IWA, so the DWA-to-IWA step is many-to-one.

We use Anthropic's bundled `onet_task_statements.csv` rather than a separately downloaded O*NET task file. Anthropic's later data documentation states that `onet_task_statements.csv` was constructed from O*NET Database 20.1, so we use the O*NET 20.1 `Tasks to DWAs` and `DWA Reference` files to keep the work-activity crosswalks on the same database version.

Step 1: Merge Anthropic task shares to O*NET task statements. Anthropic's public notebook normalizes the `Task` field in `onet_task_statements.csv` by lowercasing and trimming extraneous spaces:

$$\texttt{task\_normalized} = \texttt{Task.str.lower().str.strip()}.$$

It then merges `task_pct_v2.csv` onto `onet_task_statements.csv` using `task_name = task_normalized`. The notebook also notes that some normalized task names appear in multiple occupation-task rows, and therefore counts the number of occurrences of each normalized task and divides task prevalence across them. Specifically, the notebook computes `n_occurrences` and defines a scaled occupation-side share proportional to `pct / n_occurrences`.

Let p_t denote Anthropic's task share for normalized task t, and let n_t denote the number of matched occupation-task rows in `onet_task_statements.csv`. We assign each matched occupation-task row (o, t) the weight

$$w_{ot}^{(1)} = \frac{p_t}{n_t}.$$

We drop the Anthropic row `task_name = "none"` before proceeding. Anthropic's data documentation defines `none` as the absence of the attribute, such as no task selected, so it is not a usable O*NET task for our bridge. After dropping this row, we renormalize the remaining task mass to sum to 100.

Step 2: Map O*NET task rows to DWAs. We next merge the occupation-task rows from Step 1 to O*NET's `Tasks to DWAs` file using the pair (O*NET-SOC Code, Task ID). Because a given occupation-task row can map to multiple DWAs, a second allocation step is required. Let m_{ot} denote the number of DWA links for occupation-task row (o, t). We assign each occupation-task-DWA row (o, t, d) the weight

$$w_{otd}^{(2)} = \frac{w_{ot}^{(1)}}{m_{ot}} = \frac{p_t}{n_t m_{ot}}.$$

This allocation preserves total task mass after the task-to-DWA expansion. We then collapse to the DWA level:

$$W_d = \sum_{o,t:(o,t)\to d} w_{otd}^{(2)}.$$

Step 3: Map DWAs to IWAs. We merge the DWA-level file to O*NET's `DWA Reference` file using `DWA ID`. O*NET states that each DWA is linked to exactly one IWA. Let $i(d)$ denote the unique IWA containing DWA d. We then construct the Anthropic IWA measure by summing DWA-level mass within each IWA:

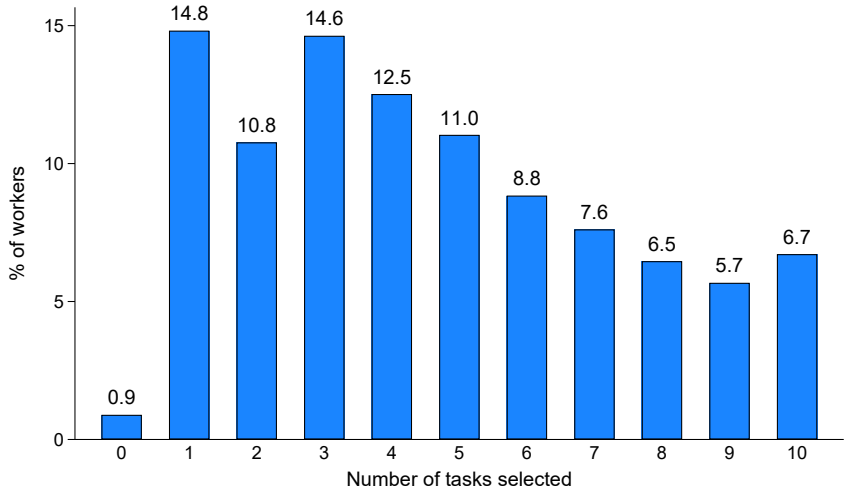
$$A_i = \sum_{d:i(d)=i} W_d = \sum_{o,t,d:i(d)=i} \frac{p_t}{n_t m_{ot}}.$$

The resulting object A_i should be interpreted as the share of Anthropic task-mapped activity assigned to IWA i after passing through the O*NET task → DWA → IWA hierarchy. It is therefore a derived IWA-level analogue of Anthropic's public task-level prevalence measure, not a measure Anthropic publishes directly.

D Figures

‘

FIGURE D.1: Number of ONET Task that Workers Report Performing

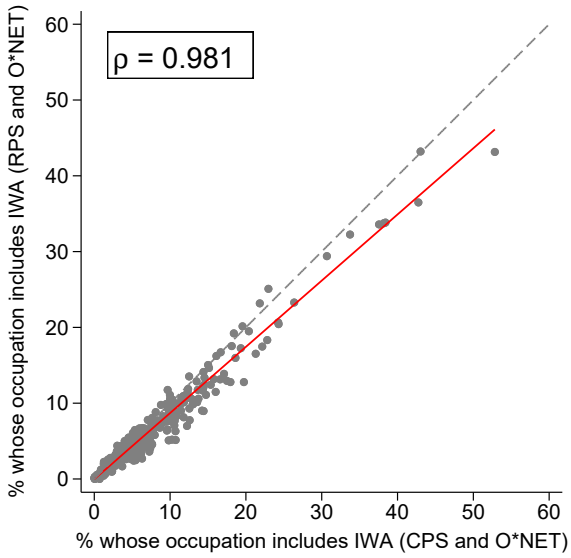


Notes: Dataset is pooled from Aug. 2025, Nov. 2025, Feb. 2026, and May 2026 RPS survey waves. Survey respondents are shown the top 10 ONET activities in their detailed occupation and asked whether they actually perform each task at work. The figure shows the distribution of the number out of the top 10 task which respondents perform.

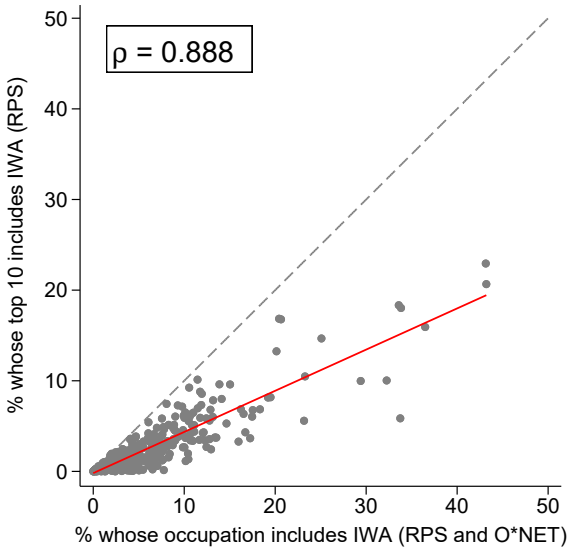
FIGURE D.2: Adoption Rates by Work Activity (WA) and Very Broad Work Activity (VBWA)

FIGURE D.3: IWA Prevalence

(a) All IWAs in Occupation - RPS vs. CPS



(b) Top 10-Truncation in the RPS



Notes: Each point is one of 332 IWAs. Panel (a): The x-axis shows the CPS employment-weighted share of occupations that include the IWA; the y-axis shows the corresponding RPS respondent-weighted share. All IWAs assigned to each occupation in O*NET are considered. Panel (b): The x-axis shows the share of RPS respondents whose occupation includes the IWA (at any rank); the y-axis shows the share whose top 10 includes it. Points below the 45-degree line indicate prevalence lost to the top-10 restriction. Dashed lined: 45-degree line. Red line: OLS fit.