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

April 27, 2026

Abstract

We measure how workers use genAI in their jobs using a nationally representative survey that links genAI use to detailed tasks. GenAI currently assists a broad range of work, with at least one in five workers using genAI in 80% of occupations and 40% of job tasks. Yet in most of these cases adoption rates remain below 50%, with some individuals systematically adopting genAI for more tasks than others who perform similar work. As a result, although genAI "exposure" measures correlate positively with adoption, they explain only about half of the variation across workers. The nature of genAI use also varies across occupations, with workers in some occupations using it mainly for high-expertise work and others for low-expertise work. Survey-based task shares differ systematically from estimates from genAI platform chat-log data, primarily because chat data overclassify basic, generic tasks relative to the ONET framework. We organize our findings into occupation- and task-level adoption indexes, providing an input for research on the labor market effects of genAI.

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 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 a joint grant from by Washington Center for Equitable Growth and the Russell Sage Foundation.

1 Introduction

Generative artificial intelligence (genAI) tools, such as ChatGPT and Claude, 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 depend critically 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 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). Since August 2024, the survey has included a genAI adoption module that asks workers whether they use genAI for their job, how intensively they use it, and how much time the technology saves them.

A key challenge in trying to learn what work genAI does is that it is infeasible to ask survey respondents about hundreds or thousands of specific work activities. We address this challenge in two steps. First, for each worker we elicit their detailed occupation using an autocomplete text box linked to the 2018 Standard Occupational Classification (SOC) system.[1] Second, in the three most recent survey waves conducted between August 2025 and February 2026, we use the worker's occupation to identify the ten most important work tasks (Detailed Work Activities) for their occupation, as described in the Bureau of Labor Statistics' ONET database. We present these ten tasks to respondents and ask them to select which tasks they performed in their job. For workers who use genAI for their job, we then ask which of these tasks they used genAI to help with. This information allows us to construct a task-level adoption rate: among workers who do a given task, what share use genAI for it?

We begin by validating our measures of tasks and task-level AI adoption rates. The median worker reports performing five tasks (out of a maximum of ten), while only 1.2% of workers select zero of the presented tasks. The share of workers who perform a given task varies substantially across tasks and is stable across survey waves, suggesting that our measure of task performance contains meaningful information. We verify that task shares in the RPS are similar to an economy-wide benchmark constructed using occupation shares and occupation-task mappings from ONET. This provides assurance that restricting attention to workers' ten

[1] Our occupation autocompletion approach follows Miano (2025). We extend his method to job titles that do not exactly match a unique SOC code or an autosuggestion by presenting respondents with a choice of probabilistic matches using the job title to a SOC code matching algorithm developed by the National Institute for Occupational Safety and Health (Laughlin et al., 2024).

most-important tasks does not present a biased view of tasks' relative importance in the US economy. Finally, conditional on performing a given task, the share of workers who report using genAI for that task also varies substantially across tasks and is stable across survey waves.

We use genAI adoption data to establish four key findings on what work genAI does. First, we characterize the current state of genAI adoption for work in the US. GenAI currently assists a broad range of work: at least one in five workers uses genAI in roughly 80% of occupations and 40% of tasks. While a few occupations show very high levels of adoption, led by computer programmers at over 90%, most occupations exhibit moderate adoption rates. For example, in roughly 40% of occupations at least a fifth of workers use genAI but more than half do not. We find similar patterns at the task level. These results suggest important individual-level determinants of adoption, which we investigate later in the paper.

The replication package for this paper will include a file containing adoption rates at the detailed task and occupation level. Summarizing at a high level, occupations with the highest genAI adoption rates are management and professional occupations, specifically in business, finance, and computer-related occupations. Tasks with the highest genAI adoption rates require cognitive and information-intensive work involving data analysis, interpretation, and writing. In contrast, the occupations with the lowest genAI adoption rates are generally personal service occupations that require manual or in-person interaction. Tasks with the lowest genAI adoption rates are concentrated in physically embodied or situational activities that require direct interaction with people or equipment. Several of these tasks also involve care or safety work, which may require quick actions and entail important responsibility or liability concerns.

Autor and Thompson (2025) argue that a key determinant of the labor market impact of a new technology is whether it primarily automates low- or high-expertise components of a job. The RPS asks genAI users whether the technology primarily assists low-, moderate-, or high-expertise work. In all occupations, a plurality of workers report that it assists with moderate-expertise work. However, we also find sizable variation across occupations. High adoption occupations with high education levels, such as Computer and Math, Architecture and Engineering, and Management have much greater rates of high-expertise assistance. Low adoption occupations with lower education levels, such as Protective Service and Building and Grounds Maintenance, have greater rates of low-expertise assistance. An interesting counterexample to this pattern is Legal occupations, who report greater rates of low-expertise assistance despite having high adoption and high education levels.

Second, we compare existing "genAI exposure" predictions with realized adoption rates across occupations and tasks. These measures use task content and technological characteristics to project 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). At both the occupation and

task level, exposure measures are predictive of actual adoption. These positive correlations between exposure predictions and realized adoption help validate that these widely used measures capture meaningful variation in which work is assisted by genAI. At the same time, exposure measures alone only explain around half of the variation in adoption across occupations and tasks, again emphasizing individual-level determinants of adoption.

Many occupations and tasks with high exposure relative to adoption relate to clerical and administrative work, which may be done by workers with modest technology skills and may involve compliance requirements or data confidentiality concerns that present barriers to adoption. Instances with high adoption relative to exposure include a variety of occupations and tasks that tend to feature high levels of autonomy, which may minimize adoption barriers.

Third, while task characteristics strongly shape how genAI is used, individual-specific factors appear to be at least as important. Within narrowly defined tasks, adoption is far from universal: even in the highest-adoption tasks a substantial share of workers do not use genAI. Consistent with this heterogeneity, adding worker fixed effects to task-level adoption regressions substantially increases explanatory power, indicating the importance of unobserved individual-level factors such as preferences, skills, beliefs, or prior experience. We provide evidence consistent with one potential driver of individual variation in genAI use: the diffusion from genAI use in one domain to another. We show that experienced users apply genAI to more tasks, workers who perform high-adoption tasks are more likely to adopt genAI for other tasks, and plausibly exogenous shifters in workplace adoption predict higher rates of home use. This evidence is consistent with a learning or experimentation mechanism: workers may incur an initial fixed cost to adopt genAI for tasks where benefits are most salient, and then expand use to a broader set of activities as they become familiar with the technology. These results underscore the value of adoption data, which reflect not only a technology's capabilities but also the individual-level choices and barriers that shape its use.

Fourth, we compare task-level use of genAI in the RPS to related measures based on genAI chat logs from Handa et al. (2025) and Chatterji et al. (2025).[2] We find systematic differences between the two chat-based measures and between our survey-based measures of genAI use at the task-level, with correlations across datasets below 0.4. These differences appear to derive from differences in how each data source assigns genAI use to tasks, which helps clarify how each measure should be interpreted.

A key distinction is that our survey-based measure begins with a worker's occupation and then asks about genAI use across that occupation's core tasks. In contrast, chat-based measures attempt to infer the most relevant task for each individual chat. We argue that chat-based

[2]In a future draft, we plan to incorporate results from a third chat-based study, Tomlinson et al. (2025), who analyze Microsoft Bing Copilot conversations.

classification procedures are predisposed to assign chats to tasks with broad, generic descriptions. For example, over 15% of chats in OpenAI data are classified as "Edit written materials or documents," even though only 1.6% of workers are in occupations that include this task in ONET. While the share of workers who perform editing is clearly much larger than 1.6%, in most occupations editing is embedded within higher-level, purpose-oriented tasks, such as preparing reports, drafting legal documents, or conducting research. More generally, the tasks most over-represented in chat data relative to the RPS all describe generic activities applicable across many contexts, such as "Resolve computer problems" or "Create visual designs or displays." As a result, task shares in chat-based measures are highly concentrated: the top five percent of tasks account for roughly 57% of all genAI use in chat data sources, compared to only 19% in the RPS.

Over-classification of generic, activity-based tasks (relative to the ONET framework) does not imply that chat-based measures are fundamentally invalid. In particular, they are likely identifying the basic work activities that genAI assists. For example, in other surveys we find that writing is the most common application of genAI among a set of basic activities, mirroring the finding in OpenAI chat data (Bick et al., 2026a). Instead, these comparisons highlight two key considerations when interpreting this data. First, chat-based occupation shares should not be interpreted literally through the ONET lens, since many chats appear to be mapped to tasks outside the user's occupation in ONET. Second, while chat-based measures are well suited to identifying the basic work activities that genAI assists, ONET mixes generic, activity-based tasks with higher-order, purpose-oriented tasks. This complicates mapping chat data into ONET and makes interpretation challenging, suggesting that alternative task frameworks that emphasize the more basic underlying activities may be more useful.

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 our survey-based measures to task-level measures constructed from genAI chat logs. Section 5 compares adoption rates to existing genAI exposure predictions. Section 6 investigates individual-level determinants of adoption. Section 7 discusses channels through which genAI may drive future productivity gains. 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 three survey waves that include these measures, conducted in 08/25, 11/25, and 02/26.

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

We elicited respondents' job titles through a free text response:

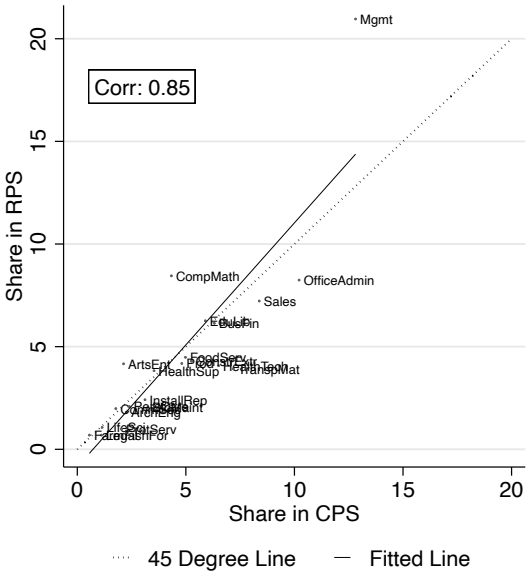
> *What is your main occupation? Please type your occupation in the box below and select one of the suggested options. Try to be specific. For instance, write preschool teacher or high school teacher rather than just teacher. If none of the options corresponds to your occupation, try adding more detail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

This follows Miano (2025). We extend his method to job titles that do not exactly match a unique SOC code or an autosuggestion by presenting respondents with a choice of probabilistic matches using the job title to SOC code matching algorithm developed by the National Institute for Occupational Safety and Health (Laughlin et al., 2024).

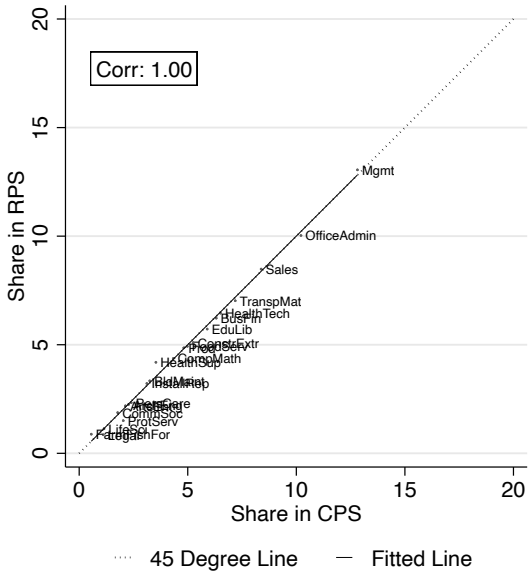
Once they have starting typing, respondents are presented with autocomplete suggestions covering over 40,000 occupations from ONET and the Occupational Outlook Handbook (OOH), following Miano (2025). We extend his method by not forcing respondents to select one of the autocomplete suggestions. Instead, 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).

Figure 1: Comparing Occupation Shares in the RPS and CPS

(a) Occupation Shares: Unweighted



(b) Occupation Shares: Weighted



Notes: The figure on the left use unweighted RPS data, the figure on the right use weighted RPS data. We use the same sample of RPS respondents in both figures. All figures use weighted CPS data. Data samples for the occupation comparison are employed respondents ages 18-64 in the pooled August 2025, November 2025, and February 2026 RPS and CPS with sample sizes of 10686 and 118784, respectively.

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 and 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.

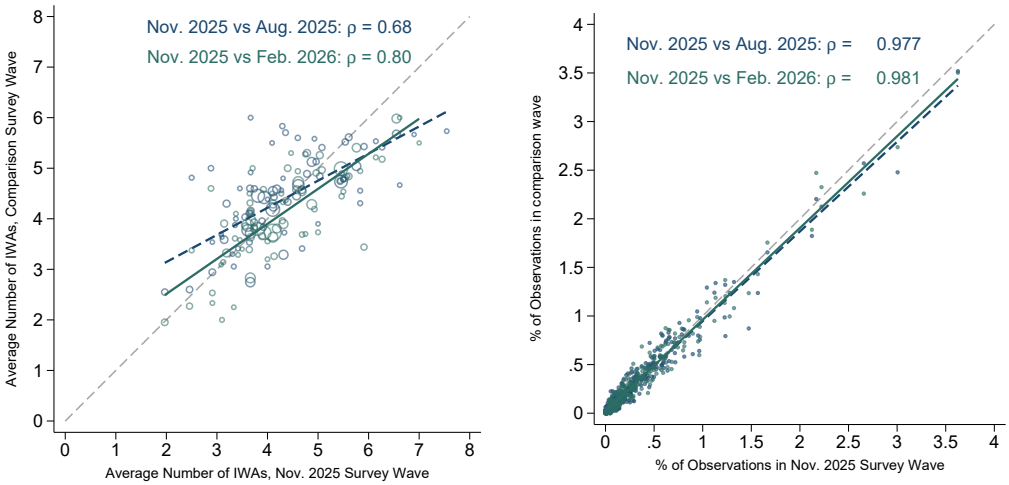
Figure 1 compares the distribution of occupation employment shares in the RPS and CPS. (Occupation shares were not targeted by our sampling scheme.) Figure 1a shows that occupation shares in the unweighted RPS data are overall very similar to those in the CPS. Our weighting scheme includes occupation 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 hundreds or thousands of detailed tasks. Our solution is to use a respondent's occupation to identify the most important tasks associated with their job and focus on these.

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. 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.

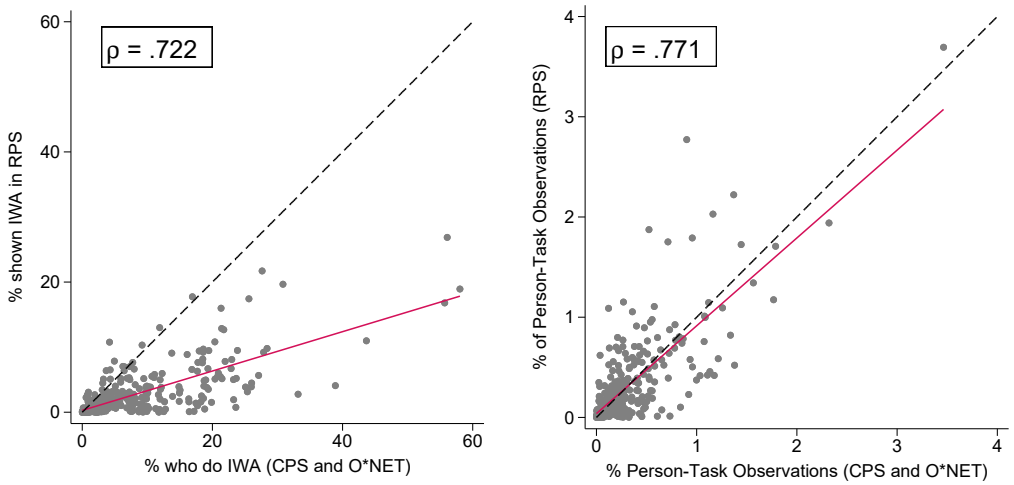
ONET contains two sets of granular activities: Tasks and Detailed Work Activities (DWAs). ONET assigns Importance Scores to each of an occupation's Tasks, which reflect how essential a Task is to overall job performance in that occupation. However, we opt to use DWAs because they tend to be more concise 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 the construction of Importance Scores, the linking procedure and precise questions 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. Later, in the genAI module, we ask respondents to select the tasks for which they regularly use genAI among the tasks they perform. 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 2a shows that the mean number of tasks selected varies by occupation. In most

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 CPS against the share of all people who, based on CPS employment shares and the associated O*NET 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 O*NET task mappings against the share of person-task observations in the truncated RPS.

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 Figure C.1.) Figure 2b shows that tasks also vary substantially in their overall prevalence. Both of these measures are stable across the three 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 in the RPS to benchmarks constructed from the CPS and ONET.

In Figure 3a, the unit of observation is a task (IWA). 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 0.722, indicating that

[3]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 C.3a. Appendix Figure C.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

task rankings are broadly similar. Moreover, because the omitted tasks are less important, the share of omitted tasks weighted by ONET importance is lower.

Figure 3b compares task shares across the two surveys. For each task t, we construct its economy-wide share of work in both datasets by counting how often the task appears across all worker–task pairs and normalizing by the total number of such pairs. In the CPS benchmark, this is based on all tasks associated with each worker's occupation in ONET, while in the RPS it is based on the subset of 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 ($\rho = 0.771$), and the best-fit line 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, although this measurement approach omits a meaningful share of work, the RPS provides a useful approximation to the economy-wide distribution of work across tasks.

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:

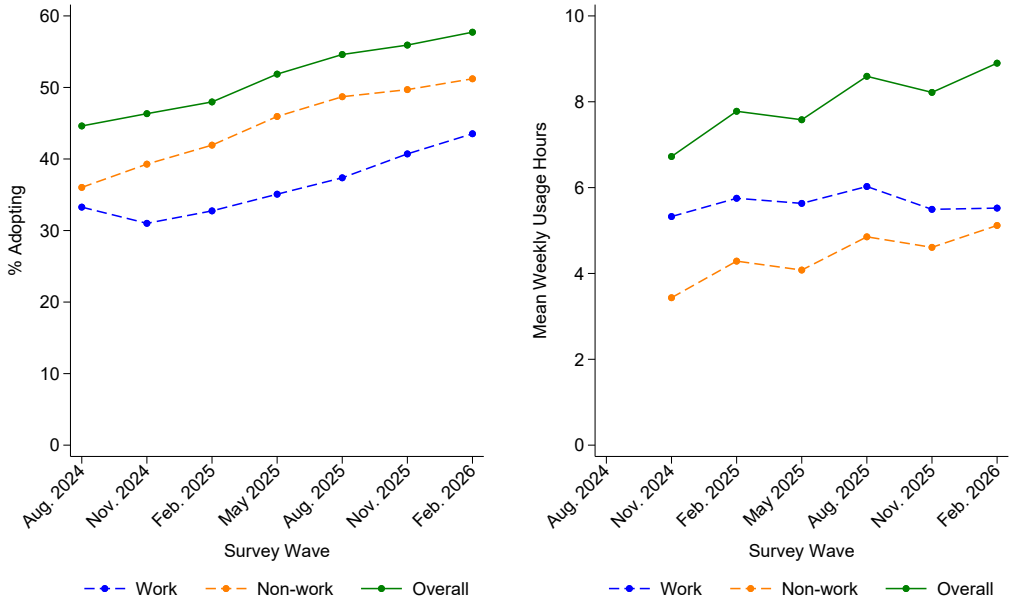
> Employed Respondents: *Do you use Generative AI for your job? (No/Yes)*
>
> All Respondents: *Do you use Generative AI [outside your job]? (No/Yes)*

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;

above the 45 degree line.

Figure 4: genAI Adoption Over Time

(a) Extensive Margin: AI Adoption Rate (b) Intensive Margin: Weekly Usage Hours



Notes: In both panels, the work use lines limit the sample to respondents who work. Panel 4a shows the share of respondents using genAI for work, non-work, and overall by RPS survey wave. Panel 4b shows the mean usage hours per person by survey wave, conditional on using genAI (and in the case of the work line, conditional on at-work use of genAI). 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.

see Appendix Section A.5. For the second question, non-employed respondents are not shown the term in brackets.

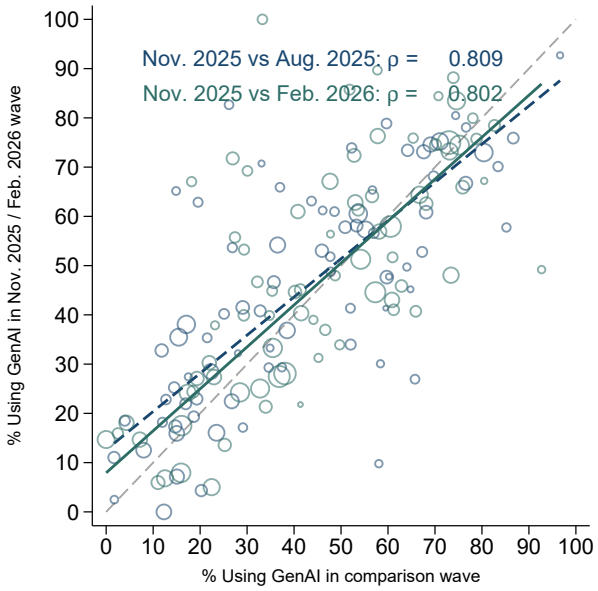
The survey asks several follow-up questions of genAI users, including which products they used, how intensively they used it, and estimates of how much time it saved them. See Bick et al. (2026b) for an overview of results.

An important caveat regarding our measurement approach is that it 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.

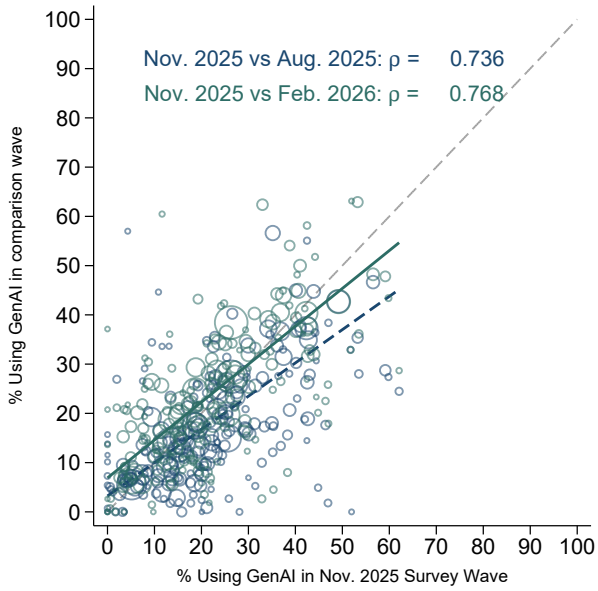
Figure 4a plots the share of individuals who report using genAI for work, non-work, and overall. As of February 2026, 51% of US adults aged 18-64 use genAI for non-work reasons, while 43% of workers use genAI for work. The overall adoption rate, i.e., the share of individuals

Figure 5: Consistency of Occupation- and Task-Level genAI Adoption Rates

(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 inside the box, ρ, is the weighted correlation coefficient between adoption weighted by number of appearances of the occupation/task in the pooled data.

who use genAI either for work or non-work purposes, is 58%. Along all three margins, adoption has increased substantially since our first reading in November 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

The RPS elicits workers' occupations and work tasks ahead of the genAI module. After asking workers whether they use genAI, 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 two waves of the RPS, in August 2025 and November 2025. Overall, the correlation coefficient is 0.809.

GenAI users are then 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 two waves of the RPS, in August 2025 and November 2025. Overall, the correlation coefficient is 0.736.

The stability of occupation- and task-level genAI adoption rates 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 Generative AI Adoption in the RPS

This section documents 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.

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, normalized to 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

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

With our normalization of $y_{o,t} = 1$, the percent change in aggregate productivity from genAI is

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

$$\omega_t \equiv \sum_o \lambda_o \omega_{o,t} \tag{4}$$

where ω_t is the total share of task t in the economy.

Within this framework, the economic impact of genAI depends on three key objects. First, we require 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, restricting attention to the top ten tasks in each worker's occupation (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.

Second, we require task-level genAI adoption rates, a_t. These adoption rates determine the extent to which potential productivity gains are realized in practice. The RPS provides direct 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 immediately below in Section 3.2.

Finally, we require task-level productivity gains from genAI conditional on adoption, g_t. The RPS contains self-reported estimates of genAI time savings, which, although not attributed to specific tasks, could be combined with information on task adoption rates to estimate the g_t. Complementary evidence on g_t can be drawn from micro-level experimental studies for selected tasks or from exposure scores.

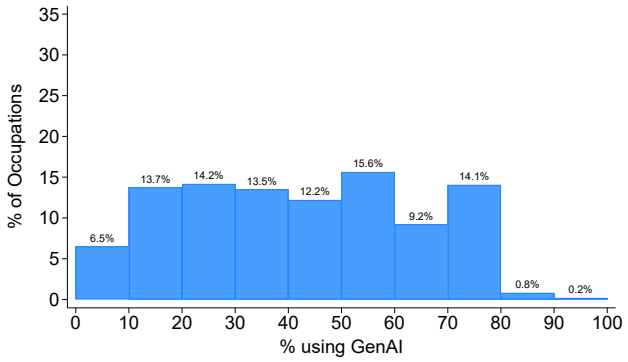
In sum, while subject to the limitations discussed above, the RPS nonetheless provides either direct estimates or empirical proxies for each of the key objects needed to bring the framework to the data.

3.2 Generative AI Adoption Rates Across Occupations and Tasks

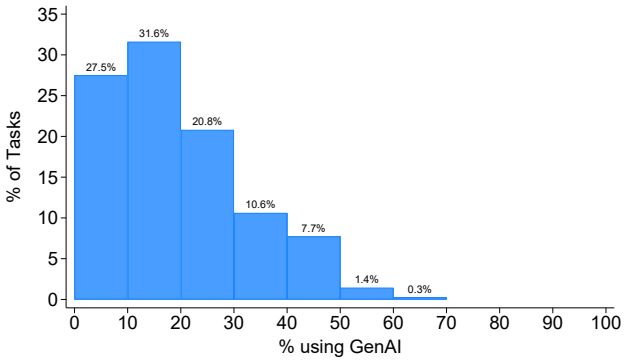
Figure 6 displays the distribution of adoption rates across detailed occupations and tasks. Together, these distributions point to three key takeaways. First, genAI adoption is already widespread and assists a broad range of work. For example, over 80% of occupations and 40% of tasks exhibit adoption rates exceeding 20%.

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 and Nov. 2025 RPS survey waves. Tasks are weighted by the number of observations in the pooled Aug. 2025 and Nov. 2025 RPS survey waves, and occupations are weighted by the number of people. Occupation is at the CPS detailed (3-digit) occupation level.

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, 39% of tasks have adoption rates between 20% and 50%. This variability within detailed occupations and tasks suggests important individual-level determinants of adoption, a point we explore further in Section 6.

Third, despite individual variation in adoption for many occupations and tasks, a subset of occupations do exhibit very high adoption. For example, roughly 15% of occupations exhibit adoption rates above 70%, with one (computer programmers) exceeding 90%. By contrast, only 2% 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.

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 is largely computer-based and 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.

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

Top 10 occupations		**Top 10 tasks**	
Occupation	**Rate (%)**	**Task**	**Rate (%)**
Computer programmers	92.3	Prepare research reports.	64.0
Public relations specialists	83.6	Read documents to gather technical information.	61.7
Network and computer systems administrators	81.0	Design integrated computer systems.	61.2
Computer, automated teller, and office machine repairers	80.9	Apply mathematical models of financial or business conditions.	61.0
Producers and directors	77.1	Develop instructional materials.	60.1
Computer and information systems managers	76.7	Design computer modeling or simulation programs.	59.5
Financial managers	76.6	Develop marketing plans or strategies.	55.9
Software developers	76.5	Manage human resources activities.	54.4
Personal financial advisors	76.4	Collaborate with others to determine design specifications or details.	54.2
Lawyers	75.3	Analyze data to identify or resolve operational problems.	54.0
Bottom 10 occupations		**Bottom 10 tasks**	
Occupation	**Rate (%)**	**Task**	**Rate (%)**
Maids and housekeeping cleaners	14.2	Collect biological specimens from patients.	0.0
Cooks	13.4	Drive trucks or other vehicles to or at work sites.	0.0
Stockers and order fillers	13.0	Inspect electrical or electronic systems for defects.	0.0
Janitors and building cleaners	9.8	Inspect work to ensure standards are met.	0.0
Driver/sales workers and truck drivers	9.4	Maintain medical equipment or instruments.	0.0
Couriers and messengers	9.0	Prepare patient treatment areas for use.	0.0
Receptionists and information clerks	8.2	Present food or beverage information or menus to customers.	0.0
Maintenance and repair workers, general	7.7	Record service or repair activities.	0.0
Fast food and counter workers	5.1	Sterilize medical equipment or instruments.	0.0
Animal caretakers	0.7	Trim trees or other vegetation.	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, and Feb. 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.

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 54–64%. 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 C.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 will be made available as part of the paper replication package.

3.3 Does Generative AI Assist High- or Low-Expertise Work?

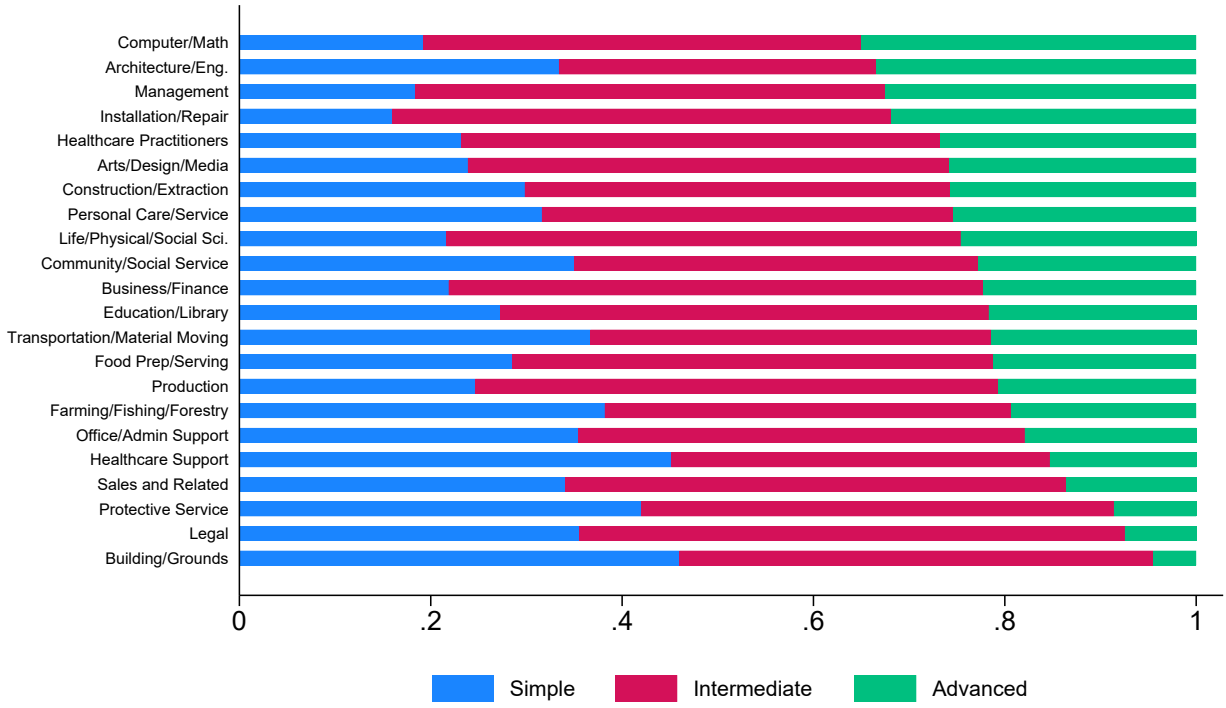
The preceding results characterize which tasks genAI assists. A complementary question is whether this assistance is primarily concentrated in simple or difficult components of a task. For example, Autor and Thompson (2025) argue that a key determinant of the labor market impact of new technologies is whether they automate low- or high-expertise work. If a technology automates high-expertise aspects of an occupation, then more workers will be able to perform the remaining (low-expertise) activities, which should increase the supply of workers in that occupation but lower average wages. Alternatively, if a technology automates low-expertise aspects of an occupation, this leaves only high-expertise work, which should lower the supply of workers in that occupation and raise average wages.

The RPS asked genAI users whether genAI primarily assisted them with low-, moderate-, or high-expertise parts of their job (workers could select multiple options). Figure 7 displays the average responses by broad occupation. In most occupations, a plurality of workers indicate that genAI assists with work that require intermediate expertise. However, we also observe substantial variation in the share of workers who report that genAI assists with low- or high-expertise work.

In most technical and professional occupations, a relatively large share of workers report that genAI assists with high-expertise work. For example, Computer and Math, Architecture and Engineering, and Management Occupations all exhibit high-expertise shares above 30%. By contrast, many occupations involving manual work exhibit small shares of workers who report high-expertise assistance. For example, Protective Service and Buildings and Grounds Maintenance occupations exhibit high-expertise shares below 10%.

However, we also observe a few interesting counterexamples to these patterns. Installation, Maintenance, and Repair Occupations, which involve substantial manual work and lower

Figure 7: Does Generative AI Assist High- or Low-Expertise Work?



Notes: Data are pooled across the August 2025, November 2025, and February 2026 waves of the RPS. The figure reports, for each CPS broad occupation, the share of respondents who use genAI and chose respective category from the following question: *"Some tasks are simpler (requiring less expertise), while others are more advanced (requiring more expertise). In your job, which types of tasks do you mainly use Generative AI for? Simple (little expertise required)/Intermediate (some expertise required)/Advanced (high expertise required)."*

average education rates, exhibit large high-expertise shares. In the other direction, Legal occupations, which exhibit high adoption rates, high education levels, and minimal manual work, exhibit small high-expertise shares.

Overall, these patterns suggest that genAI does not map cleanly into a uniform "low-expertise" or "high-expertise" case. In most occupations, a plurality of workers indicate that genAI assists with work that requires some, but not high, expertise. Moreover, occupations vary considerably in their share of high-expertise assistance. Through the lens of Autor and Thompson (2025), these results suggest that the labor market implications of genAI are likely to vary considerably across occupations, even among those with similar adoption rates.

4 Comparing GenAI Adoption and Chat Data

The RPS measures genAI use via workers' self-reports. An alternative approach starts from genAI platform chat logs and classifies chats into specific tasks. We compare our survey-based measures to those of Handa et al. (2025) and Chatterji et al. (2025) below; in a future draft we we plan to incorporate results from Tomlinson et al. (2025). By comparing survey- and chat-based estimates, we can learn about the strengths and limitations of each. Before turning to the empirical comparison, we first show how chat-based measures map into the conceptual framework introduced in Section 3.1.

4.1 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 $\sum_t c_t = 1$.

To map these objects into the economic framework in Section 3.1, consider first an idealized setting in which (i) the task shares ω_t and adoption rates a_t of the platform's users match the broader economy, (ii) every genAI-assisted task on the platform entails the same number of chats, n, and (iii) chats are classified across tasks in a manner perfectly consistent with ONET. Then we can write

$$C_t = \omega_t a_t n \tag{5}$$

indicating that the number of chats corresponding to task t is the product of the task's share of work in the economy, the adoption rate for task t, and the chat-intensity n.

A trivial implication of the above equation is that, on their own, chat data will not be informative about the level of adoption rates. Intuitively, because chat data only observe instances in which genAI is used, they are uninformative about work in which genAI is not used and, as a result, the share of work assisted by genAI. However, Equation (3) makes clear that this information is essential for understanding the economic impact of genAI.[4]

[4]In our framework, g_t is the productivity gain from genAI on task t conditional on adoption. If instead g_t were recast as the productivity gain from a single genAI chat, chat data could then be informative about productivity gains. Nearly all existing estimates of the productivity gains from genAI are based on comparisons of task or job performance with and without genAI use, rather than at the level of individual chats. An interesting example attempting to estimate productivity gains from individual chats is Tamkin and McCrory (2025).

Define the overall task adoption rate by

$$a \equiv \sum \omega_t a_t \tag{6}$$

This implies that the relationship between chat shares and adoption rates is

$$a_t = \frac{c_t}{\omega_t} a \tag{7}$$

This equation is simply Bayes' Rule: the probability of adoption in task t conditional on performing t (a_t) is the probability of performing task t conditional on adoption (c_t), multiplied by the overall probability of adoption (a), divided by the overall probability of task t (ω_t).

Should we expect chat shares to match relative adoption rates across tasks? The above equation implies that the ratio of chat shares for tasks s and t is

$$\frac{c_t}{c_s} = \frac{a_t \omega_t}{a_s \omega_s}. \tag{8}$$

This equation shows that chat shares will not generally reflect relative adoption rates because chat shares conflate adoption with task prevalence. The ratio of chat shares for two tasks c_t and c_s equal the ratio of adoption rates a_t and a_s only when the two tasks are equally common in the economy, $\omega_t = \omega_s$. Because task shares vary widely (see Figure 5b), chat shares and adoption rates will typically differ. Recovering adoption rates from chat shares therefore requires an external measure of ω_t. One potential proxy for ω_t are ONET task importance scores combined with occupational employment shares. 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.

4.1.1 Practical Measurement and Comparison Challenges

The above analysis relied on several assumptions which may not hold in practice.

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 B for details). We highlight three considerations based on these definitions. First, comparisons of task shares between OpenAI and Anthropic data could be influenced by different definitions of a chat. Second, especially for Anthropic data, a single conversation may

actually assist multiple distinct tasks if users do not uniformly start a new conversation when alternating between tasks. Third, the scope for task misclassification may differ depending on the level of the task hierarchy at which chats are initially classified.

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 most important 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.

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.

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

Variation in Chat Intensity The analysis above assumes that each instance of genAI-assisted work generates the same number of chats, n. In practice, chat intensity is likely to vary across tasks, for two reasons. First, some tasks occupy a larger share of work time than others. The economic framework and RPS data do not include this intensive margin. Second, genAI assistance in some tasks require many iterative chats, while assistance in other tasks may require only a single prompt. (This latter aspect is likely more relevant for the OpenAI analysis, which defines a chat as a single message.) Let n_t denote the average number of chats per genAI-assisted instance of task t. Then chat counts satisfy $C_t = \omega_t a_t n_t$, so that chat shares reflect both adoption rates and differences in chat intensity across tasks. As a result, variation in n_t introduces an additional wedge between chat shares and measured adoption rates.

4.2 Comparing Empirical Estimates

Figure 8 compares genAI task shares, c_t for three sources of task-level data: the RPS and two chat-based data sets from OpenAI and Anthropic. For the RPS, c_t is the share of task t among all AI-assisted worker–task pairs:

$$c_t = \frac{a_t}{a}\omega_t \tag{9}$$

For OpenAI and Anthropic data, c_t is the share of chat messages classified to task t. All three are vectors of shares summing to one.

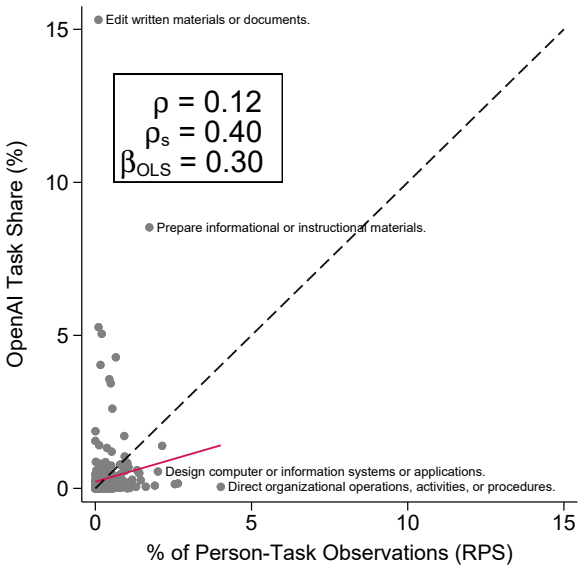
An immediate takeaway is that all three measures display substantially different task shares from one another. The correlation between the RPS and chat data sources are 0.12 (OpenAI) and 0.34 (Anthropic). The correlation between the two chat data sources is 0.38.

A second takeaway is that the RPS features fewer tasks with very large task shares. The task with the highest share in each chat data source exceeds 15%, compared to roughly 4% in the RPS. More broadly, in OpenAI data, nine tasks (out of 332 total IWAs) have shares exceeding 2%, with a total share of roughly 52% among them. Similarly, in Anthropic data ten tasks have shares exceeding 2%, with a total share of roughly 44%. By contrast, in the RPS only three tasks have shares exceeding 2%, with a total share of roughly 9%. Figure 9 confirms that the distribution of genAI task shares is much more concentrated in the chat data than in the RPS. For example, ranked by task shares, the top 5% of tasks account for roughly 57% of genAI use in chat data, versus 19% of genAI use in RPS data.

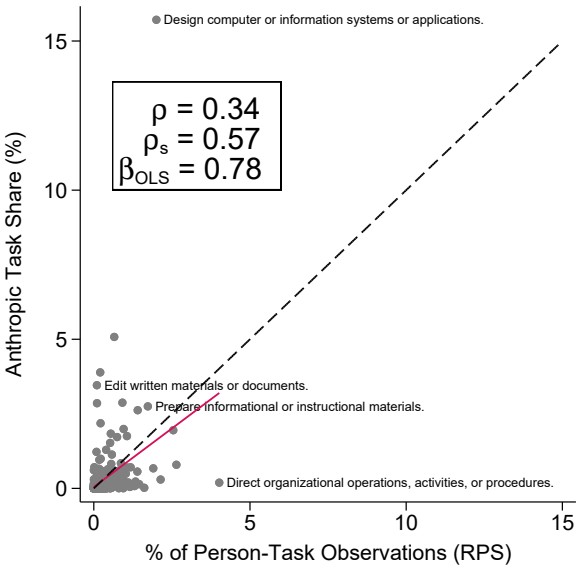
To investigate the sources of disagreement between data sources, Table 2a summarizes the share of the sum of squared differences accounted for by the largest outliers. Panel (a) shows that between 47-60% of the total (squared) differences between the RPS and the chat data sources is driven by a single task. For OpenAI, this is "Edit written materials or documents."

Figure 8: Comparing GenAI Task Shares in RPS, OpenAI, and Anthropic Data

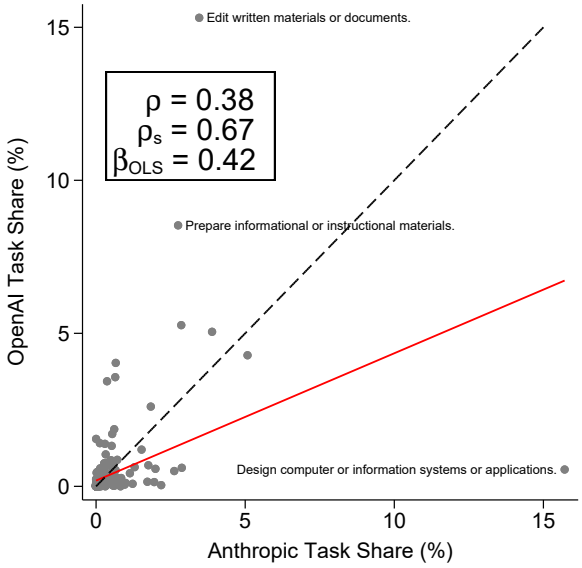
(a) RPS vs. OpenAI Shares



(b) RPS vs. Anthropic Shares

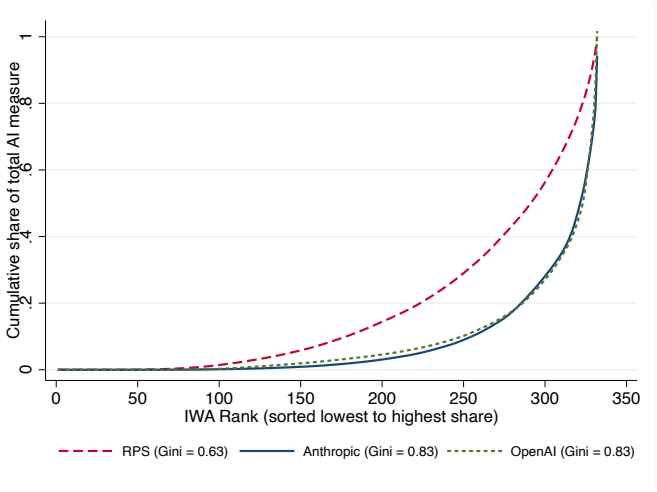


(c) OpenAI vs. Anthropic Shares



Notes: Data for the RPS percent of person-task observations are pooled across the August 2025, November 2025, and February 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 slope of OLS regression (β_{OLS}) for the displayed sample.

Figure 9: Cumulative Distribution of IWA Task Shares



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

For Anthropic, this is "Design computer or information systems or applications." Roughly 90% of the total differences are driven by only ten tasks.

The tasks driving this disagreement are informative. Panel (a) of Table 3 lists the ten tasks most overrepresented in chat data relative to the RPS (averaging across OpenAI and Anthropic). These titles all describe basic, generic actions that apply across many purposes and contexts—for example, "Edit written materials or documents," "Resolve computer problems," and "Create visual designs or displays." Notably, these tasks all exhibit high adoption rates in the RPS: that is, low genAI task shares in the RPS do not imply low adoption rates conditional on performing the task. Instead, the RPS reports lower shares because relatively few workers are employed in occupations for which these tasks are listed in ONET.

Panel (b) of Table 3 shows the opposite pattern: the tasks most underrepresented in chat data tend to be higher-order, purpose-oriented titles. For example, "Develop organizational policies, systems, or processes" encompasses many more basic actions such as brainstorming, analysis, writing, and communicating with others.

As described in Section 4.1.1, one possible interpretation of these results is that chat classification procedures may be predisposed to assigning chats to tasks whose titles reflect basic, generic actions. As a concrete and quantitatively important example, only 1.5% of workers are in occupations that include the IWA "Edit written materials or documents." This does not imply that only 1.5% of workers edit written materials or documents: this must be more than an order of magnitude larger. Instead, most occupations which involve editing written materials or documents include other higher-order, purpose-oriented tasks which encompass editing and writing as subcomponents. When a genAI user requests editing help, however, the

Table 2: GenAI Task Share Disagreements Are Concentrated in a Small Number of Tasks

(a) All Tasks

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	OpenAI vs. Anthropic
1	50.9%	60.3%	47.2%
5	75.5%	79.2%	87.2%
10	87.1%	86.3%	93.1%
$\sum_t d_t^2$	0.0455	0.0312	0.0487

(b) Tasks Overrepresented in Chat Data

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	OpenAI vs. Anthropic
1	59.0%	71.6%	62.7%
5	87.3%	91.4%	90.7%
10	97.3%	96.0%	96.5%
$\sum_t d_t^2$	0.0393	0.0262	0.0224
# IWAs	97	98	164

(c) Tasks Underrepresented in Chat Data

Top k IWAs	OpenAI vs. RPS	Anthropic vs. RPS	OpenAI vs. Anthropic
1	25.3%	29.6%	87.5%
5	53.7%	52.1%	94.2%
10	65.6%	64.9%	97.1%
$\sum_t d_t^2$	0.0062	0.0049	0.0263
# IWAs	215	218	133

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.

basic and generic title may appear to be the most literal fit for the chat content. If so, these IWAs could serve as catch-all categories that absorb genAI interactions from many different occupational contexts. This interpretation is consistent with the asymmetry documented in Table 2. 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 a larger set of tasks. (By contrast, disagreements between the two chat data sources find similarly concentrated differences for both under- and over-represented tasks.) This pattern aligns closely with the mechanism described above: a small number of generic, activity-based task labels may act as catch-all categories that absorb chats from many different contexts, which results in the under-representation of a large number of higher-order, purpose-oriented tasks.

A related question is whether these disagreements also reflect fine-grained classification

Table 3: The Ten Most Over- and Underrepresented Tasks

IWA	Shares in % RPS	Chat avg	pp diff	% RPS Adoption
Panel A: Chat overrepresents (top 10)				
Edit written materials or documents.	0.1	9.4	9.3	49.6
Program computer systems or production equipment.	0.2	4.5	4.3	32.7
Resolve computer problems.	0.7	4.7	4.0	39.8
Write material for artistic or commercial purposes.	0.1	4.1	4.0	32.2
Prepare informational or instructional materials.	1.7	5.6	3.9	37.3
Develop marketing or promotional materials.	0.2	2.3	2.2	31.4
Create visual designs or displays.	0.5	2.2	1.7	34.2
Gather information from physical or electronic sources.	0.4	2.1	1.7	31.4
Interpret language, cultural, or religious information for others.	0.0	1.2	1.2	4.3
Analyze scientific or applied data using mathematical principles.	0.5	1.4	0.8	42.9
Panel B: Chat underrepresents (top 10)				
Direct organizational operations, activities, or procedures.	4.0	0.1	-3.9	31.1
Maintain operational records.	2.6	0.5	-2.2	20.1
Supervise personnel activities.	1.6	0.0	-1.6	17.9
Respond to customer problems or inquiries.	1.9	0.4	-1.5	15.9
Analyze data to improve operations.	2.5	1.0	-1.5	39.2
Communicate with others about operational plans or activities.	2.1	0.8	-1.3	20.1
Develop organizational policies, systems, or processes.	1.5	0.2	-1.3	38.5
Execute financial transactions.	1.3	0.1	-1.2	11.7
Resolve personnel or operational problems.	1.2	0.1	-1.1	43.4
Monitor individual behavior or performance.	1.1	0.1	-1.0	15.9

Notes: IWAs ranked by average difference between the two chat-log measures and the RPS, restricted to IWAs where both OpenAI and Anthropic agree on the sign of the difference relative to the RPS. "Chat avg" is the average of the OpenAI and Anthropic task shares. "diff" is (Chat avg − RPS), expressed in percentage points. Panel A lists the 10 IWAs most overrepresented in chat logs; Panel B lists the 10 most underrepresented.

noise among similar tasks that washes out at higher levels of aggregation, or more fundamental differences in what the data sources capture. Table 4 reports pairwise correlations between task shares at three levels of the ONET hierarchy: Intermediate Work Activities (IWA, 332 categories), Work Activities (WA, 37), and Broad Work Activities (BWA, 9). Between the two chat data sources, the correlations increase sharply at higher levels of aggregation: the Pearson correlations increase from 0.38 at the IWA level to 0.72 at the WA level and 0.94 at the BWA level. This suggests that fine-grained classification noise between similar tasks can potentially explain a large share of the disagreement between OpenAI and Anthropic task shares.

Between the RPS and either chat source, the correlations also increase at higher levels of aggregation, but to a smaller degree. Pearson correlations climb from roughly 0.12–0.34 at the IWA level to 0.54–0.58 at the WA level and 0.56-0.58 at the BWA level. This suggests that fine-grained classification differences likely explain some but not most of the disagreement between chat data and the RPS: even when tasks are grouped into just nine broad categories, the two types of data sources continue to tell meaningfully different stories.

Table 4: GenAI Task Share Correlations For Various Aggregation Levels

	N	OpenAI vs. RPS	Anthropic vs. RPS	OpenAI vs. Anthropic
Correlation				
IWA	332	0.12	0.34	0.38
WA	37	0.54	0.55	0.72
BWA	9	0.58	0.56	0.94
Rank Correlation				
IWA	332	0.40	0.57	0.67
WA	37	0.52	0.77	0.76
BWA	9	0.68	0.70	0.83

Notes: Pearson correlations and Spearman rank 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). N denotes the number of task categories at each level.

Summary Taken together, the evidence in this section points to systematic differences between chat-based and survey-based measures of genAI use. A central source of disagreement relates to 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 survey-based measure maps tasks into tasks tied to occupations by ONET, which include higher-order, purpose-oriented tasks that are underrepresented in chat data. This difference leads to highly concentrated chat shares and more diffuse survey-based measures, with a small number of generic tasks accounting for a large share of the divergence across survey- and chat-based data sources. Differences in task shares between chat sources likely appear to primarily reflect fine-grained differences in classification procedures, which cancel out at higher levels of task aggregation. Differences in user populations and variation in chat intensity across tasks may also contribute to these discrepancies, though we are not able quantify the importance of these factors.

Over-classification of generic, activity-based tasks (relative to the ONET framework) does not imply that chat-based measures are fundamentally invalid. Instead, these results highlight three 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. Second, chat-based occupation shares should not be interpreted literally, since these measures are constructed using ONET but many chats appear to be mapped to tasks that ONET states are outside the user's occupation. Third, these considerations suggest 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.

5 Comparing GenAI Adoption and Exposure Predictions

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.

5.1 Mapping Exposure Scores Into the Conceptual Framework

In our conceptual framework, task exposure scores proxy for predicted task-level productivity gains, g_t. Exposure scores do not directly measure the magnitude of productivity gains, but instead predict the tasks in which g_t may be relatively large. For example, Eloundou et al. (2024) construct 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 a set of task-level exposure score e as an indicator function which equals one if a task's predicted productivity gain exceeds a threshold $\overline{g}_e$:

$$e_t = \mathbb{1}\{g_t \geq \overline{g}_e\} \tag{10}$$

In the case of Eloundou et al. (2024), $\overline{g}_e = 1$, i.e., the threshold corresponds to 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 based on task similarity to machine learning capabilities. 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.

5.1.1 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

$$a_t = \Pr(g_t \geq \chi_{i,t} | i \text{ performs } t) = F_t(g_t). \tag{11}$$

The above equation implies that, if an exposure score is accurate for two tasks s and t with the same adoption cost distribution, $F_s = F_t$, the more exposed task will have the higher adoption rate.

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

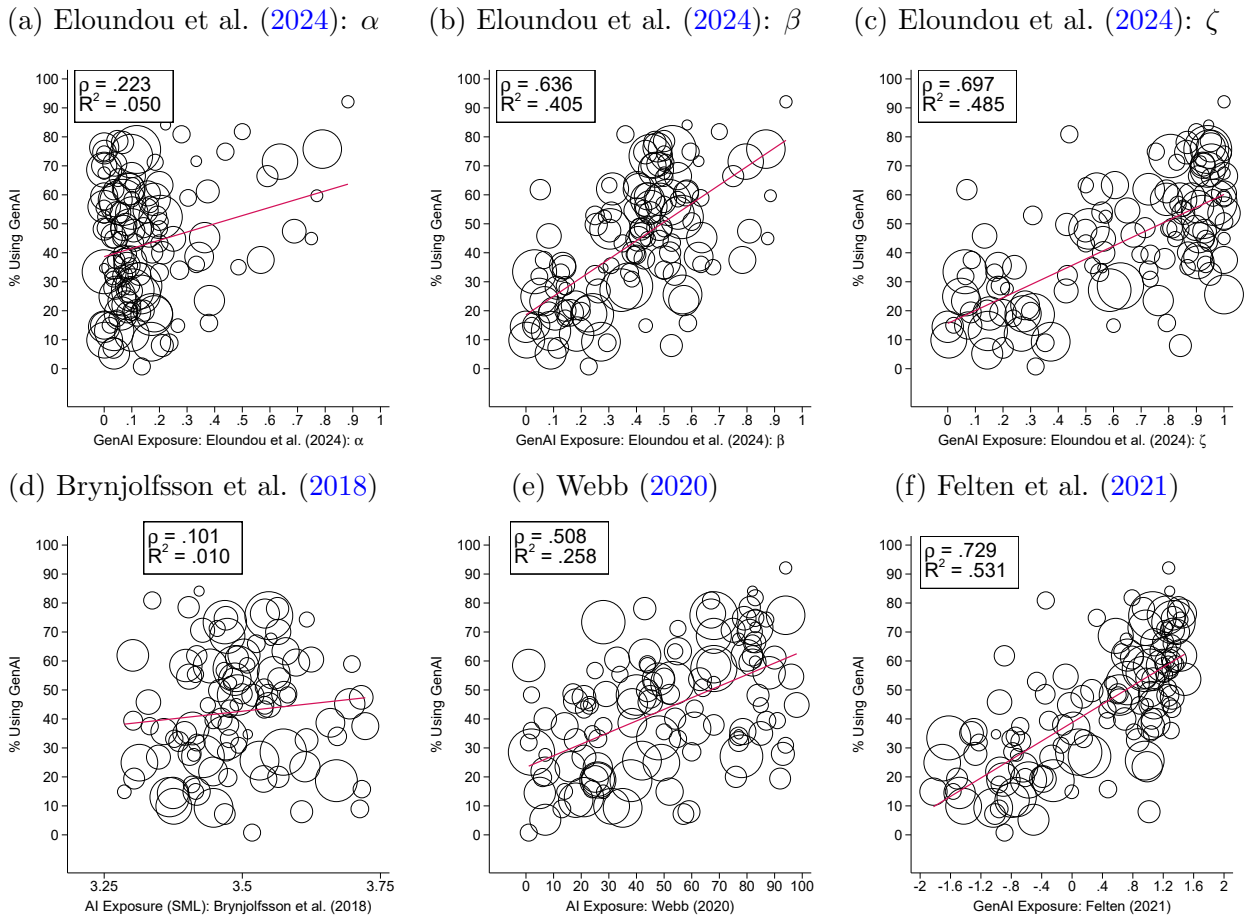
$$\chi_{i,t} = \mu_t + \xi_i + \epsilon_{i,t} \tag{12}$$

Task-specific barriers μ_t reflect barriers that are common to all workers. For example, some tasks may require substantial 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.

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 scores 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.

Figure 10: GenAI Adoption versus Exposure Predictions by Occupation



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, and Feb. 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.

5.2 Do Exposure Scores Predict GenAI Adoption?

Figure 10 compares occupation-level genAI adoption rates from the RPS to six genAI exposure scores. Figures 10a–10c 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 are the measures from Webb (2020) and Felten et al. (2021).

[5]We use the original GPT-4-based scores throughout. Yin et al. (2026) show that the Eloundou et al. (2024) methodology produces substantially different exposure scores when applied to more recent models.

Table 5: Predicting GenAI Adoption

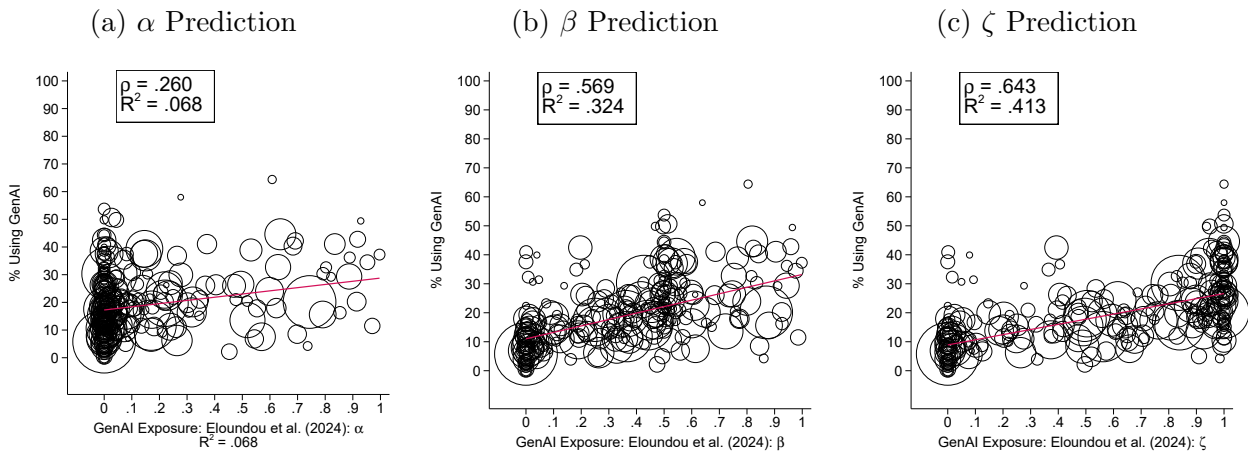
Controls	Adjusted R^2	
	Worker Level	Task Level
Eloundou et al. (2024) ζ Exposure Measure	0.077	0.025
Felten et al. (2021) Exposure Measure	0.084	
Detailed Occupation Fixed Effects	0.184	0.121
DWA Fixed Effects		0.108

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, and Feb. 2026 survey waves, and all regressions include a fixed effect for survey wave.

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 substantial gaps between predicted exposure and actual adoption to date. Column 1 of Table 5 summarizes how well different occupation-level measures predict whether an individual worker actually uses genAI for their job. We compare three occupation-level measures: (i) the ζ exposure exposure measure from Eloundou et al. (2024); (ii) the Felten et al. (2021) exposure measure; and (iii) average adoption rates by occupation in the RPS. All three measures are predictive of worker-level adoption, but their explanatory power differs substantially. Occupation-level adoption rates yield an adjusted R^2 of 0.184, compared to 0.077 for Eloundou et al. (2024) and 0.084 for Felten et al. (2021). That is, occupation-level adoption rates more than double the share of explained variation relative to occupation-level exposure scores.

The occupation-level exposure measures in Eloundou et al. (2024) are constructed from task-level predictions, which allows us to directly compare their task-level exposure scores to observed genAI adoption in the RPS. Figure 11 presents these results. Consistent with the occupation-level patterns, the α measure is only weakly correlated with task-level adoption, while the β and ζ measures, which assume greater genAI capabilities, are more strongly correlated. As with the occupation-level results, this pattern suggests that exposure scores are informative about which tasks are actually assisted by genAI. However, the correlations between adoption and the β and ζ exposure measures are lower at the task level than at the occupation level, indicating somewhat larger gaps between exposure predictions and realized adoption at the task level. Column 2 of Table 5 repeats the prediction exercise at the worker-task level, where the dependent variable is an indicator for whether the worker uses genAI for a given task. The Eloundou et al. (2024) ζ exposure measure yields an adjusted R^2 of 0.025, well below its

Figure 11: 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, and Feb. 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.

worker-level counterpart. DWA fixed effects yield an adjusted R^2 of 0.108, more than four times the explanatory power of exposure scores at the task level. Detailed occupation fixed effects yield a slightly higher R^2 of 0.121, indicating that even for the same task, adoption rates differ systematically across occupations. As with the worker-level results, adoption-based measures substantially outperform exposure-based measures in predicting realized genAI use.

5.3 Where Do Exposure and Adoption Diverge?

Which occupations and tasks exhibit large gaps between adoption and exposure? For simplicity, we focus on comparisons between adoption in the RPS and the ζ exposure scores from Eloundou et al. (2024). We regress task- and occupation-level adoption rates on exposure scores, which yields a predicted adoption rate for each task. Table 6 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: receptionists, customer service representatives, tellers, office clerks, and bookkeepers. 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, reflecting modest technology skills, or weak incentives to adopt. Second, task-level frictions μ_t may be high in these roles: in particular, standardized workflows, compliance requirements, and confidential data limitations may prevent widespread adoption of genAI. The most over-predicted tasks include prescribing medical treatments, examining

Table 6: 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	8.0	53.1	−45.1	Computer, ATM, and office machine repairers	80.9	35.2	45.7
Tellers	15.8	50.9	−35.1	Construction equipment operators	61.8	18.7	43.1
Customer service representatives	25.6	60.1	−34.5	Computer programmers	92.1	60.1	32.0
Animal caretakers	0.8	29.8	−29.0	Public relations specialists	84.1	57.7	26.4
Musicians and singers	14.9	42.3	−27.5	Network and computer systems administrators	81.8	55.7	26.1
Office clerks, general	23.5	49.5	−26.0	Computer support specialists	74.9	49.1	25.7
Sales representatives, wholesale/manufacturing	32.7	57.9	−25.2	Special education teachers	63.3	37.9	25.5
Driver/sales workers and truck drivers	9.3	32.2	−22.9	Automotive service technicians and mechanics	45.8	21.6	24.3
Bookkeeping, accounting, and auditing clerks	37.5	60.1	−22.6	Producers and directors	78.1	54.2	23.9
Couriers and messengers	9.0	31.3	−22.4	Shuttle drivers and chauffeurs	52.9	29.3	23.6

Panel B: Tasks

Most over-predicted tasks				Most under-predicted tasks			
Task	**Act.**	**Pred.**	**Gap**	**Task**	**Act.**	**Pred.**	**Gap**
Operate communications equipment or systems.	4.2	26.4	−22.1	Develop models of systems, processes, or products.	64.4	26.6	37.8
Prescribe medical treatments or devices.	5.0	25.0	−20.0	Test sites or materials for environmental hazards.	40.9	8.9	32.0
Examine materials/docs for accuracy or compliance.	8.5	26.5	−18.0	Evaluate project feasibility.	58.0	26.6	31.3
Assist others to access services or resources.	10.5	26.6	−16.1	Obtain formal documentation or authorization.	39.9	10.3	29.6
Operate office equipment.	2.2	17.7	−15.5	Signal others to coordinate work activities.	37.6	8.9	28.7
Prepare health or medical documents.	11.5	26.6	−15.1	Develop business or marketing plans.	53.7	26.6	27.1
Estimate project development or operating costs.	11.6	26.6	−15.1	Resolve personnel or operational problems.	42.5	15.9	26.6
Evaluate patient/client condition or treatment options.	4.8	18.1	−13.3	Investigate organizational or operational problems.	49.8	25.5	24.3
Notify others of emergencies or problems.	7.6	20.6	−13.0	Analyze performance of systems or equipment.	49.8	25.7	24.1
Schedule appointments.	13.7	26.6	−13.0	Manage human resources activities.	50.7	26.6	24.0

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. "Actual" is observed genAI adoption, "Pred." is the fitted value based on exposure, and "Gap" is actual minus predicted adoption. Rows are ordered by the absolute value of the residual. All statistics are based on pooled Aug. 2025 and Nov. 2025 survey waves.

documents for accuracy or compliance, and preparing medical documents, which may prohibit genAI use for regulatory or privacy reasons. Over-predicted tasks also include operating office equipment and processing shipments or mail, which could potentially be aided by genAI but might require substantial complementary capital investments and workplace reorganization.

Under-predicted occupations include a variety of occupation types, including machine operators and repairers, drivers, and computer programmers. 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 training. Under-predicted tasks include developing models, analyzing data and risk, developing business plans, which genAI appears well-suited to assist given the requisite knowledge of what needs to be done. Under-predicted tasks also include two tasks, "Apply hygienic or cosmetic agents to skin or hair" and "resolve personnel or operational problems" which have very low predicted exposure scores. One potential interpretation is that these tasks, which involve physical or interpersonal work that genAI may not be well-suited to address, can still benefit from generic feedback that genAI can provide: for example, the specific steps to apply a cosmetic product, or general advice for resolving personnel disputes.

6 Learning and the Diffusion of GenAI Across Domains

We have shown that genAI adoption rates vary substantially across tasks. At the same time, in many cases adoption is highly variable across individuals performing the same task. For example, Figure 6 showed that in roughly 40% of tasks, at least a fifth of workers use genAI while more than half do not. What explains this within-task variation?

Table 7 regresses task-level genAI adoption on task fixed effects and various individual-level controls. Observations are at the worker-task level, which allows us to explore both task and individual fixed effects. Task fixed effects alone generate an adjusted R^2 of 0.07, indicating that 7% of the variation in adoption occurs between tasks. This reflects the fact that adoption rates vary substantially across tasks, and yet in most tasks adoption is not uniformly zero or one.

One potential explanation for within-task variation in adoption is that, for some reason, individuals vary in their propensity to adopt genAI. This could reflect differences in demographic characteristics, such as sex, age, race, and education. However, the second row of Table 7 shows that controlling for these characteristics only marginally increases the share of adoption variation explained.

The third row of Table 7 shows that including individual fixed effects dramatically in-

Table 7: Predicting Adoption: Individuals vs. Tasks

Controls	Adjusted R^2
Task Fixed Effects	0.074
Task Fixed Effects + Demographic Controls	0.098
Task Fixed Effects + Individual Fixed Effects	0.426

Notes: Tasks are at the Intermediate Work Activity (IWA) 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 46,438 person-task observations. Demographic Controls consist of education, race, sex, and a quadratic for age.

creases the adjusted R^2 to 0.435. The sharp increase implies that a large share of variation in task-level genAI adoption reflects persistent individual differences among workers with similar demographics. In other words, for a given task, some workers are systematically more likely to adopt genAI than others. This suggests that individual characteristics, such as skills, beliefs, learning costs, willingness to experiment, or the individual's work environment, play a central role in adoption decisions.

Our data do not allow us to investigate most of these characteristics directly. What they do contain, however, is information on worker's 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: prior experience with genAI. Prior genAI experience may teach workers about genAI's relative strengths and limitations, how to effectively prompt genAI, and how to integrate genAI into pre-existing workflows. Crucially, this knowledge may translate from one domain to another: for example, once a worker has experimented with genAI for one task, they may not have to experiment again to use genAI effectively in the next task.

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 empirically investigate it.

6.1 Conceptual Framework: Learning as a Cost Shifter

Recall from Section 3.1 that worker i adopts genAI for task t when the productivity gain exceeds the adoption cost:

$$a_{i,t} = \mathbf{1}\{g_t \geq \chi_{i,t} = \mu_t + \xi_i + \varepsilon_{i,t}\}. \tag{13}$$

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 learning how to use genAI through experience, we let ξ_i itself depend on the worker's accumulated genAI experience:

$$\xi_i = \bar{\xi}_i - \theta \cdot E_i, \tag{14}$$

where E_i measures prior genAI experience, $\bar{\xi}_i$ is a fixed underlying propensity, 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 workers with low $\bar{\xi}_i$ are more likely to adopt genAI and therefore accumulate experience, so a raw correlation between adoption and proxies for experience would likely reflect both 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. No single test is decisive, but similar findings across distinct analyses makes the learning interpretation more credible than any single analysis on its own.

6.2 Three Empirical Tests of the Learning Channel

Past GenAI Use and the Breadth of Current Adoption Our first test exploits variation in how long workers have been using genAI. In the framework, E_i is proxied by a binary indicator for whether the worker was already using genAI six months ago. 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, tenure is uncorrelated with $\bar{\xi}_i$ after conditioning on occupation, demographics, and the total number of tasks performed. This is an admittedly strong assumption, but we view it as a natural starting point before proceeding to more rigorous tests.

Table 8 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

Table 8: Prior GenAI Experience and Task-Level Adoption

	Number of genAI-Assisted Tasks
Have used genAI 6 months ago	0.421***
	(0.124)
Adjusted R^2	0.461
Observations	5,058

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

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.

Cross-task spillovers Our second test asks whether workers are more likely to perform a given task if they use genAI in other tasks. We begin by constructing, for each individual-task, 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 $\bar{\xi}_i$. 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

Table 9: Other-Task genAI Use and Focal-Task Adoption

Occupation Controls	(1) None	(2) Broad	(3) Detailed
Panel A. OLS: Dependent variable = Uses genAI for focal task			
Other-task AI use share	0.632***	0.629***	0.597***
	(0.013)	(0.013)	(0.013)
Observations	38734	38734	38734
Panel B. 2SLS: Other-task AI use share instrumented by OAI			
Other-task AI use share	0.762***	0.741***	0.808***
	(0.051)	(0.054)	(0.054)
First-stage F-stat.	255.34	177.08	173.96
Observations	38734	38734	38734
Panel C. First stage: Dependent variable = Other-task AI use share			
Other-task adoption index (OAI)	0.767***	0.731***	0.751***
	(0.048)	(0.055)	(0.057)
Observations	38734	38734	38734

Notes: 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-wave-IWA observations, with IWA-level genAI use equal to one if the worker uses genAI for any DWA within that IWA. 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 occupation controls are omitted, included at the broad level, or included at the detailed level. 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.

according to their underlying adoption propensity, $\bar{\xi}_i$, which is unlikely to hold in practice. Alternatively, if we use detailed (3-digit) occupation controls, these controls absorb more of the variation in the instrument. For transparency, we display three specifications which alternately exclude occupations or control for broad (2-digit) or detailed (3-digit) occupations.

Table 9 reports the results. The 2SLS coefficients range from 0.74 - 0.81 (first-stage F scores ranging from 173 - 255), indicating that a 10 percentage-point increase adoption for a worker's other tasks is associated with a 7.4 - 8.1 percentage point increase in the worker's adoption for the focal-task. The OLS estimate estimates range from 0.59 - 0.63, slightly smaller than the 2SLS, suggesting that contamination from intrinsic-propensity sorting is modest.[6]

[6]The slightly lower OLS estimate than the 2SLS estimates is consistent with classical measurement error

Table 10: 2SLS Regression of Non-Work Usage Onto Work Usage

	Uses genAI for Non-Work Purposes
Uses genAI for Work	0.533***
	(.045)
Observations	3,815
Adjusted R^2	0.210

Notes: Regression is at the worker level, and instruments for work usage with employer encouragement of work usage. Regression includes controls for detailed occupation, industry, sex, education, college major, and survey wave fixed effects. Regression includes a regression constant. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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 home adoption is proxied by work adoption (Bick et al., 2026b). Because work adoption may itself reflect $\bar{\xi}_i$, we instrument for it using whether the respondent's employer encourages genAI use. The identifying assumption is that employer encouragement affects home-domain adoption only through its effect on work-domain adoption, conditional on detailed occupation, industry, education, college major, demographics, and survey wave. This rules out, for example, that tech-savvy workers sort into tech-friendly employers in ways that also directly affect their home adoption.

Table 10 shows a strong positive relationship: workers encouraged to use genAI at work are significantly more likely to use genAI at home. One interpretation is that plausibly exogenous exposure to genAI in the workplace induces the worker to pay the fixed cost of adoption, after which they apply the technology more broadly in their non-work activities. This result reinforces the broader pattern in this section: genAI adoption in one domain is strongly predictive of adoption in another.

attenuating OLS since "Other-task AI use share" is an average over a small number of binary indicators per worker.

7 Potential Channels of Future GenAI Productivity Gains

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

$$\frac{Y(\tau) - Y_0}{Y_0} = \sum_t \omega_t(\tau) \cdot a_t(\tau) \cdot g_t(\tau). \tag{15}$$

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.

Task reallocation Economy-wide task shares $\omega_t(\tau)$ can change over time as well (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.

GenAI improvements Even tasks with high adoption rates could still experience additional

productivity gains from genAI if genAI capabilities, $g_t(\tau)$, improve.

8 Conclusion

This paper measures how workers use genAI for their jobs by linking survey-based adoption data to detailed ONET tasks for a nationally representative sample of U.S. workers. We present four main findings.

First, genAI adoption is already widespread: over 80% of occupations and 40% of tasks exhibit adoption rates above 20%. High-adoption work is concentrated in cognitive and information-intensive activities such as analysis, writing, and evaluation, while physical, manual, and in-person work exhibits low adoption. A plurality of workers report that genAI assists with moderate-expertise components of their work, though the share reporting high-expertise assistance varies substantially across occupations.

Second, adoption rates are correlated with existing measures of genAI exposure, but exposure measures account for at most 53% of the variation in adoption across workers. Adoption-based predictors at the task- or occupation-level substantially outperform exposure measures in predicting individual workers' genAI use.

Third, adoption patterns reflect not only task characteristics but also individual-level factors. Worker fixed effects explain far more variation in task-level adoption than task characteristics alone, and we document patterns consistent with a learning mechanism: experienced users apply genAI to more tasks, workers who perform high-adoption tasks are more likely to adopt genAI for other tasks, and plausibly exogenous shifters in workplace adoption predict home use.

Finally, task-level genAI use measured in the RPS differs systematically from task shares inferred from genAI platform chat logs. Chat-based measures tend to concentrate usage in a small set of generic, activity-based tasks (e.g., editing, programming), while survey-based measures are more diffuse and tied to higher-order, purpose-oriented tasks from ONET. These differences reflect how each data source assigns chats or responses to tasks and highlight essential considerations when comparing results from different types of data sources.

A key contribution of this paper is the construction of task- and occupation-level genAI adoption indexes that can be incorporated into existing analyses of productivity, wages, task reallocation, and inequality. Unlike exposure measures, which predict genAI's potential effects, our indexes reflect how workers are actually using the technology in practice. Our estimates can be updated as genAI capabilities and adoption patterns evolve, providing an evolving input for research on the labor market effects of this new technology.

References

Autor, D. and N. Thompson (2025). "Expertise". In: *Journal of the European Economic Association* 23.4, pp. 1203–1271.

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., 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.

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.

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.

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].

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.

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., 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. Table A.1 compare the sample composition between the pooled August 2025, November 2025, and February 2026 RPS and the CPS along the demographics targeted in the sampling procedure for our main survey. The

[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 August 2025, November 2025, and February 2026 CPS and RPS

	Everyone		*Employed*	
	CPS	RPS	CPS	RPS
	(1)	(2)	(3)	(4)
Gender: Women	50.5	52.1	47.4	48.4
Age				
18-24	15.1	12.5	11.9	12.7
25-34	22.5	21.0	24.1	22.5
35-44	22.4	26.4	24.8	28.0
45-54	19.9	23.2	21.6	22.6
55-64	20.2	17.0	17.7	14.2
Race/Ethnicity				
Non-hispanic White	54.9	57.6	56.3	58.2
Non-hispanic Black	12.5	15.4	11.6	15.2
Hispanic	21.3	20.2	21.1	19.9
Other	11.2	6.8	11.0	6.6
Education				
Highschool or less	36.1	30.8	32.0	26.8
Some college/Associate's degree	26.0	27.7	25.4	27.1
Bachelor's or Graduate degree	37.8	41.5	42.6	46.1
Marital Status: Married	50.0	50.5	52.9	53.4
Number of children				
0	60.6	55.8	59.7	53.2
1	17.7	19.1	18.0	20.3
2	13.8	15.7	14.6	17.0
3+	7.9	9.4	7.8	9.6
Household Income in Last 12 Months				
$0-$50,000	24.1	31.5	18.3	24.7
$50,000-$100,000	29.5	27.2	30.1	28.9
$100,000-$150,000	18.7	19.0	20.4	21.4
$150,000+	27.8	22.3	31.2	25.0
Region				
Northeast	17.1	19.0	17.2	19.9
Midwest	20.1	16.9	20.9	16.6
South	38.8	39.9	38.0	39.3
West	24.0	24.2	24.0	24.2
Employment Status				
Employed, at work last week	71.9	75.2		
Employed, absent from work last week	2.3	3.2		
Unemployed	3.5	7.4		
Not in the labor force	22.3	14.2		
Observations	160232	14849	118784	11637

Notes: Column 1 reports the sample composition in the pooled August 2025, November 2025, and February 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, and February 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).

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.

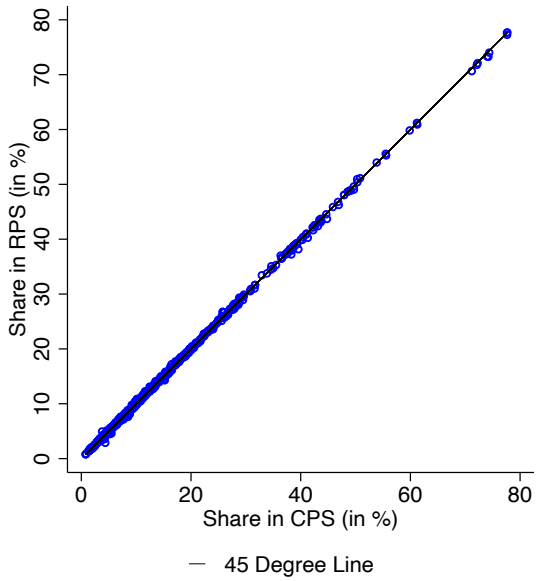
Using the iterative proportional fitting (raking) algorithm of Deming and Stephan (1940), we construct sampling weights to ensure the RPS matches the CPS sample proportions for the same set of demographic characteristics included in the Qualtrics sampling targets for the overall sample, i.e., independent 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 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 and November 2025. 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.

FIGURE A.1: Weighted Sample Composition in the Pooled August 2025, November 2025, and February 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).

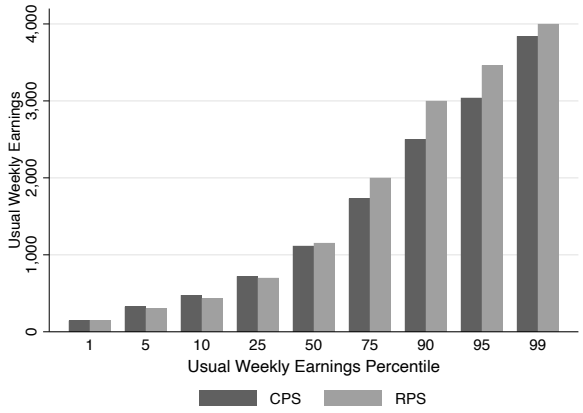
A.3 Validation Checks

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, and February 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.

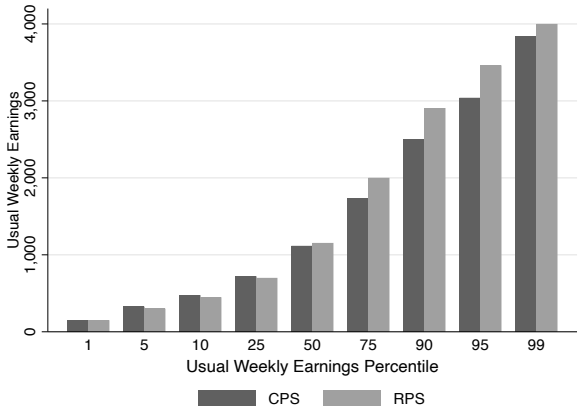
Figure A.3 compare the industry shares in the RPS and CPS in the pooled August 2025, November 2025, and February 2026 waves, with and without weights, as well as college major shares in the RPS and 2022 American Community Survey (ACS). The unweighted correlations of the industry shares is 0.74. The correlation of college majors in the RPS and the ACS is 0.75. 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.

FIGURE A.2: Validation Checks – Usual Weekly Earnings

(a) Weekly Earnings Percentiles: Unweighted



(b) Weekly Earnings Percentiles: Weighted

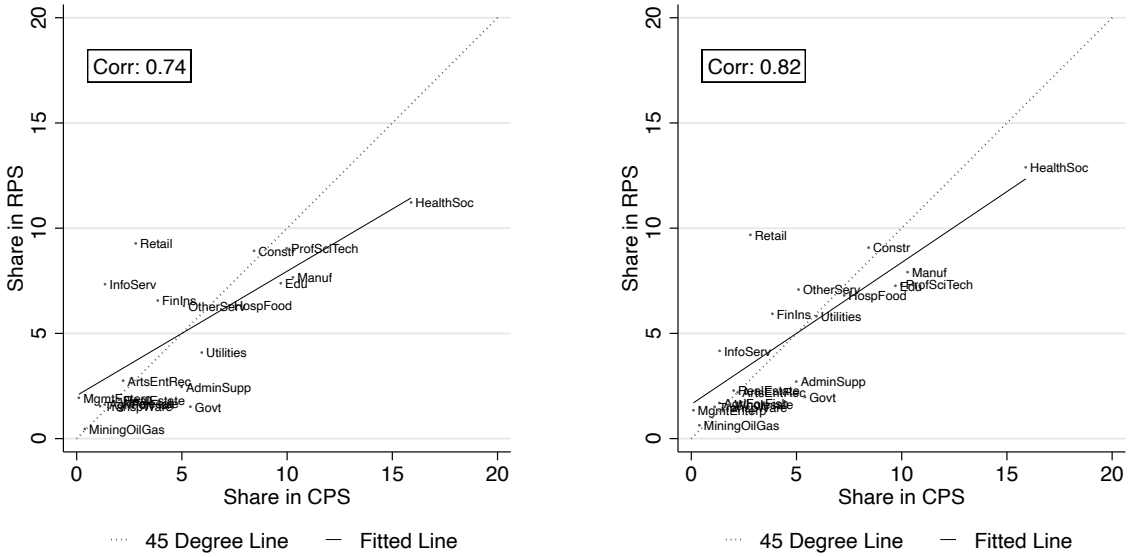


Notes: The figure on the left uses unweighted RPS data, the figure on the right uses weighted RPS data. We use the same sample of RPS respondents in both figures. All figures use weighted CPS data. Data samples for the weekly earnings figures are employees ages 18-64 in the pooled August 2025, November 2025, and February 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 both months.

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

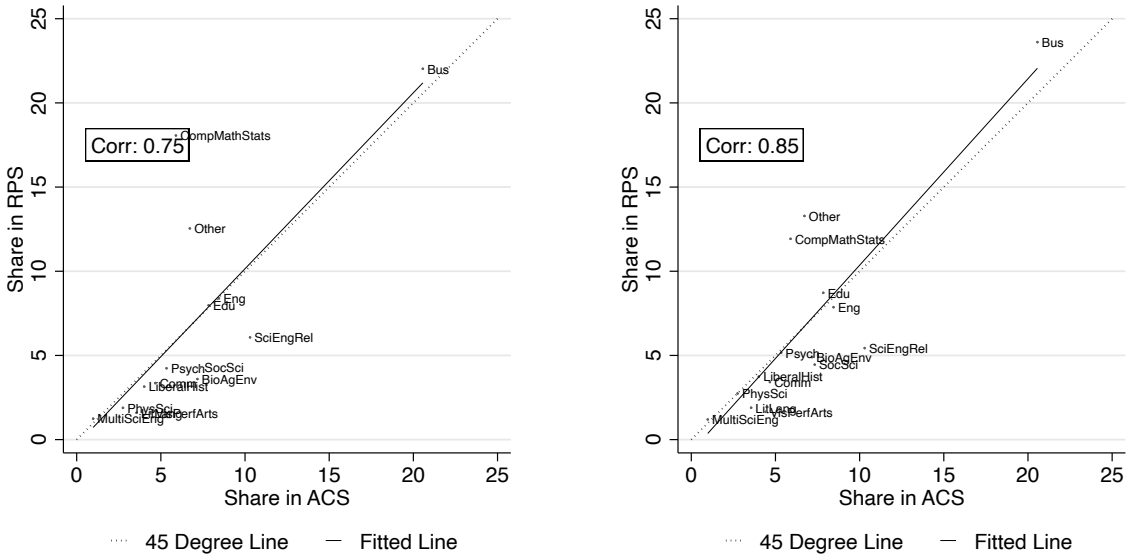
(a) Industry: Unweighted

(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, and February 2026 RPS and CPS with sample sizes of 10686 and 118784, respectively. Data samples for the college major comparison are respondents with a college degree ages 18-64 in the pooled August 2025, November 2025, and February 2026 RPS with sample sizes of 5779 and 118784 in the 2022 ACS, respectively.

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 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:

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

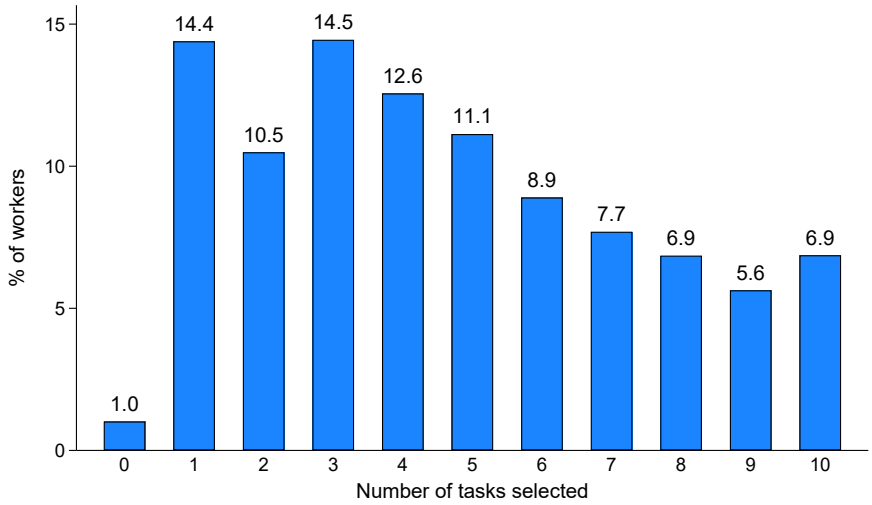
Step 3: Map DWAs to IWAs. We merge the DWA-level file to O*NET's `DWA Reference` file using `DWA ID`. O*NET states that each DWA is linked to exactly one IWA. Let $i(d)$ denote the unique IWA containing DWA d. We then construct the Anthropic IWA measure by summing DWA-level mass within each IWA:

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

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 $\rightarrow$ DWA $\rightarrow$ IWA hierarchy. It is therefore a derived IWA-level analogue of Anthropic's public task-level prevalence measure, not a measure Anthropic publishes directly.

C Figures

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

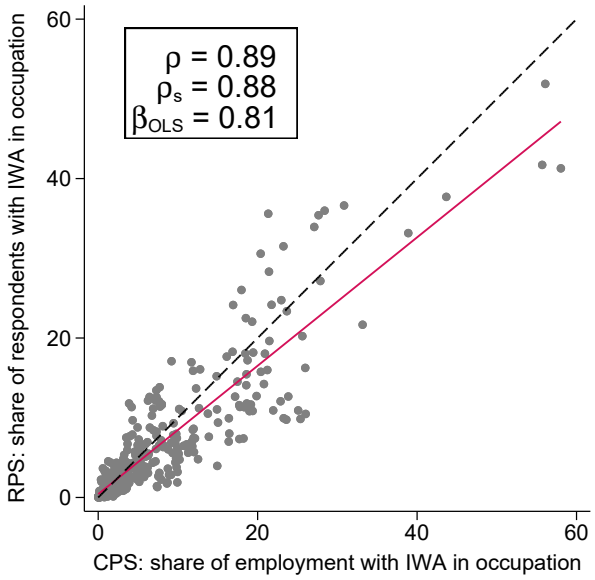


Notes: Dataset is pooled from Aug. 2025, Nov. 2025, and Feb. 2026 RPS survey waves. Survey respondents are shown the top 10 ONET activities in their detailed occupation and asked whether they actually perform each task at work. The figure shows the distribution of the number out of the top 10 task which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

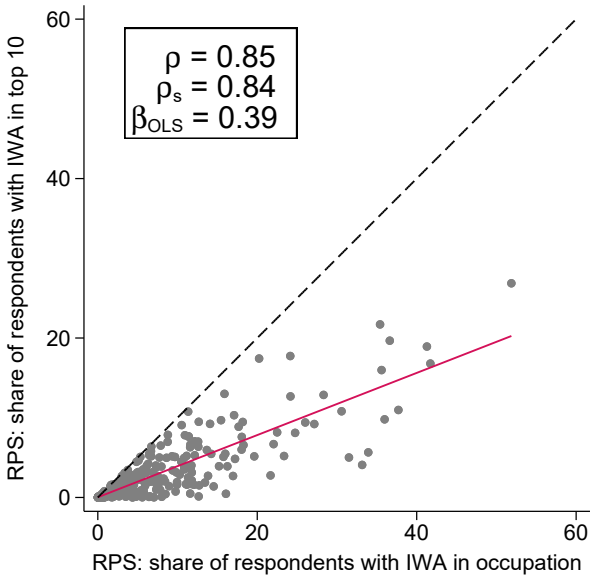
FIGURE C.2: Adoption Rates by Work Activity (WA) and Very Broad Work Activity (VBWA)

FIGURE C.3: IWA Prevalence

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



(b) Top 10-Trucnation 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.