# AI and Coder Employment: Compiling the Evidence

Leland D. Crane*  Paul E. Soto*

March 20, 2026

**Abstract**

We evaluate whether LLMs have had any discernible impact on the aggregate labor market so far. We focus on occupations that are computer programming-intensive, motivated by data showing that coding is one of the most LLM-exposed tasks. Linking O*NET to CPS we find that aggregate employment of coders has decelerated sharply since the introduction of ChatGPT. Using a novel control variable for industry-level shocks we show that the deceleration is not attributable to the exposure of coders to slowing industries, suggesting instead that coders experienced an occupation-specific shock around the introduction of ChatGPT. Coder employment has continued to grow in recent years, though much more slowly than it did pre-2022. We validate the industry-level control variable by examining historical examples of occupations that experienced either occupation-specific or industry-level shocks. We also provide statistics on the agreement rates between different measures of AI exposure.

---

*Board of Governors of the Federal Reserve System

We thank Gabriella Galassi, Dave Byrne, conference participants at the San Francisco Fed and the Bank of Canada, and brown bag participants at the Federal Reserve Board and Sogang University for insightful comments. We thank Nicole Hoffmann for excellent research assistance.

# 1 Introduction

Recent progress in the field of AI has fueled speculation that labor market disruptions may be imminent or already underway. Bloomberg asserts that "The AI Hiring Pause Is Officially Here", and prominent technology commentators speculate that fundamental transformation may be 1-2 years away.[1] Such a scenario would be a serious challenge for policy makers; in this paper we bring the available data to bear on the question.

Recently released real-world AI usage data shows that computer programming is one of the predominant uses of generative AI. The Anthropic Economic Index (AEI, Handa et al. 2025) publishes anonymized data on how people use Anthropic's flagship chatbot, Claude, a major competitor to OpenAI's ChatGPT. This is consistent with anecdotal evidence that coding assistants are one of the more successful applications of the tools. Motivated by this—and other evidence developed below—we focus on coding-intensive occupations. Our view is that if generative AI is to substantially affect the job market, the effects should be apparent here first.

We use Occupational Information Network (O*NET) data to identify occupations where computer programming is an important skill. We track monthly employment of "coders" in the CPS. Our analysis is essentially an event study; we explore whether employment in coding-intensive jobs had a markedly different trend after the introduction of ChatGPT in November 2022.[2] Simple regressions show that the growth of coder employment did indeed slow down post-ChatGPT. This is more clear when we control for industry-level shocks (discussed below.) Controlling for factors that affect industry employment but not its composition, we find robust evidence that annual coder employment growth is about 3 percent lower now than it was pre-ChatGPT. The interpretation of this estimate is subject to several caveats: among other issues, we are not controlling for AI's effects on aggregate

[1] See "Behind the Curtain: A white-collar bloodbath", "Situational Awareness" and AI 2027 for examples.

[2] Of course, the introduction of LLMs for coding is not a single event. GPT 3.5 and Github Copilot were available before November 2022, and subsequent model releases have improved coding proficiency since then. But November 2022 is the most natural breakpoint, and likely the time when most managers and business leaders would have become aware of the possible productivity implications.

labor demand or prices, occupational task mix may be changing in a way that overstates the effects on coders as a skill group, and there may be ample room for coders to match well in other occupations. Nonetheless, the results suggest that AI is having a measurable and potentially consequential impact on employment for some groups.

An important step we take is to develop counterfactual employment series for coders—that is, coder employment absent an occupational-specific shock caused by LLMs—based on industry exposure. The counterfactual is derived from a within-industry/between-industry decomposition of occupational employment growth: We explore which employment fluctuations are attributable to industry-level factors and which are occupation-specific shocks. The counterfactual variable is the sum of industry employment growth rates weighted by the distribution of coders across industries. This is the between-industry component of occupation growth; it tells us what coder employment would have done if it stayed a constant fraction of employment within each industry as those industries grew and shrank. The intuition and key identifying assumption is that industry-level shocks (like demand for the output product and changes in industry TFP) should scale industry employment up and down homothetically—not affecting composition—while occupation-specific shocks will change the composition of employment within industries. Comparing the counterfactual to actual coder employment within the information sector shows that some of the slowing in coder employment is attributable to industry-level dynamics rather than an occupation-specific shock, though a substantial fraction is still attributed to occupation-specific factors.[3]

Aside from the main results on coder employment growth and the properties of the counterfactual, we uncover several other interesting facts. For example, while coders are widely dispersed across industries, about 40 percent of coders work in computer systems design and related services (NAICS 5415). This industry covers many software/IT contractor activities. It is remarkable that the modal coder is not working at a Silicon Valley tech firm, nor at a start up, nor as an in-house developer in other industries, but is doing contract

[3]Note that the counterfactual has the form of a Bartik instrument, but we are not using it to obtain causal variation quantities; here it plays the role of a control variable.

software development. Note that this industry is not in the information sector (NAICS 51) often used to proxy the tech industry.

In addition, we explicitly compare the generative AI exposure measures from Eloundou et al. (2024) and Handa et al. (2025). The large and growing literature studying generative AI largely relies on these measures but little has been done to explicitly compare them. We find notable disagreement in the measures, though both agree that coders are among the most highly exposed occupations.

Before turning to the data and results it is important to clarify the scope of the paper. An important question is what we should expect LLMs to do to coder employment. LLM-powered coding assistants appear to be complements in production to coders: The coding assistant increases the marginal product of a coder by letting them complete tasks faster. If the demand for coding services is inelastic this could lead to a fall in employment, as fewer coders are required to satiate demand. On the other hand, elastic demand for coding services could result in increased coder employment as the more efficient coders are able to serve a much larger market for low-cost coding. This effect seems more likely in the long run than in the short run, as businesses have time to adapt and develop new products. Ex ante it is not clear which world we live in, though there is some work on the topic (Humlum, 2021; Acemoglu and Loebbing, 2026). In the longer run other important dimensions are the introduction of new work and the reorganization of occupational tasks (Acemoglu and Restrepo, 2019). LLMs may lead to the introduction of new coding-intensive products, and individuals from other occupations may begin to do LLM-assisted coding tasks. This means that both the short-run and long-run effects of AI on coder employment are empirical questions.

Relatedly, our focus is on generative AI as an occupation-specific shock. Our emphasis on exposed occupations more or less forces us to ignore other margins, such as the automation of a wide range of a firm's business processes by LLM-using coders. Our view is that initial effects of AI are most likely to show up as occupation-specific shocks, since larger-scale automation and new businesses take time to develop. While these margins are likely to dominate in the long run they might not be informative this early in the diffusion process. We also cannot address general equilibrium effects, where automation raises productivity and labor demand for all workers. These are important and difficult questions that are beyond the scope of this paper.

## 2 Literature

While future effects on productivity are uncertain, generative AI is showing signs of being a general purpose technology (GPT) and an invention of the method of invention (IMI), which could have longer lasting impacts on productivity growth (Baily et al., 2025). AI adoption rates in the workplace have also been rising steadily (Bonney et al. 2024, Bick et al. 2024, Crane et al. 2025), and studies have found that many jobs are exposed to generative AI (see Eloundou et al. 2024 and Felten et al. 2023). These findings have raised questions about the macroeconomic impact of generative AI, see, e.g., Acemoglu (2025) and Korinek and Suh (2024).

Focusing on the labor market, Humlum and Vestergaard (2025) find little imprint of generative AI use on worker wages and employment, while Brynjolfsson et al. (2025) finds a decline in the employment of young workers relative to older workers within the occupations most exposed to AI. Our work is largely complementary in that we seek to identify changes in *total occupation employment* for a large, highly-exposed occupation, while Brynjolfsson et al. (2025) identify relative changes in the age composition of employment for a broader group of exposed occupations. Brynjolfsson et al. (2025) are able include firm-level controls from their proprietary data which is not possible in the public CPS we use, though we develop useful industry-level controls. Our view is that the first-order question is whether labor demand (in a given occupation grouping) is increasing or decreasing as a result of AI. A larger literature is addressing this and related topics, including Lichtinger and Hosseini Maasoum (2025), Eckhardt and Goldschlag (2025), Gimbel et al. (2025), Iscenko and Millet (2026), Brynjolfsson et al. (2026), Dominski and Lee (2025), Atkinson and Yamco (2026), Ahn and Carollo (2026), and Massenkoff and McCrory (2026)

The literature has largely focused on indirect generative AI exposure measures developed by researchers: Eloundou et al. (2024) and Felten et al. (2023) calculate exposure by asking humans (or LLMs) to judge AI exposure of individual tasks. Soto (2025) estimates firm-level AI usage through conference call transcripts. These methodologies are useful and especially valuable when tools are new, but we now have some real-world data on usage from the Anthropic Economic Index (Handa et al. 2025). These data are our starting point. We go further than much of the literature in comparing the Handa et al. (2025) exposure metric to Eloundou et al. (2024) and showing that while there is substantial disagreement, both agree that coder occupations are highly exposed.[4]

Similar to Chandar (2025) our methodology provides timely readings of employment trends related to AI based on public data. Another contribution of our paper is to develop a control for industry-level shocks which is especially useful when firm-level data are not available. The industry-level control lets us cleanly separate occupation-specific shocks (which is our focus) from possibly correlated cyclical or secular industry factors. The control is based on simple within-between decompositions like those used by Davis and Haltiwanger (1992), Jaimovich and Siu (2020) and others, though our application is—as far as we know—original.

A separate literature (e.g., Eisfeldt et al. (2023)) examines generative AI through the lens of financial markets. They find that markets expect firms with more-exposed workforces to deliver more value in the future. Similarly, Wiles and Horton (2025) find that while AI technology can reduce private search costs for firm hiring, there is little evidence to suggest that this leads to improved labor market efficiency.

Tomlinson et al. (2025) also analyze real-world LLM use, in their case the Bing Copilot tool. We suspect that the usage patterns they document may be dominated by the fact that

[4]See Cottier et al. (2023) for an analysis of the differences between various exposure measures, prior to the release of the AEI data, and Gimbel et al. (2025) for a comparison more similar to ours.

the tool is integrated into a search engine, so many queries are of the type "looking for information." Nonetheless, our methods could be used with their data as well.

## 3 Motivation and Coders as an Occupational Group

In this section we first provide evidence that coders are likely the most generative AI-exposed occupational group and then construct a definition of coders.

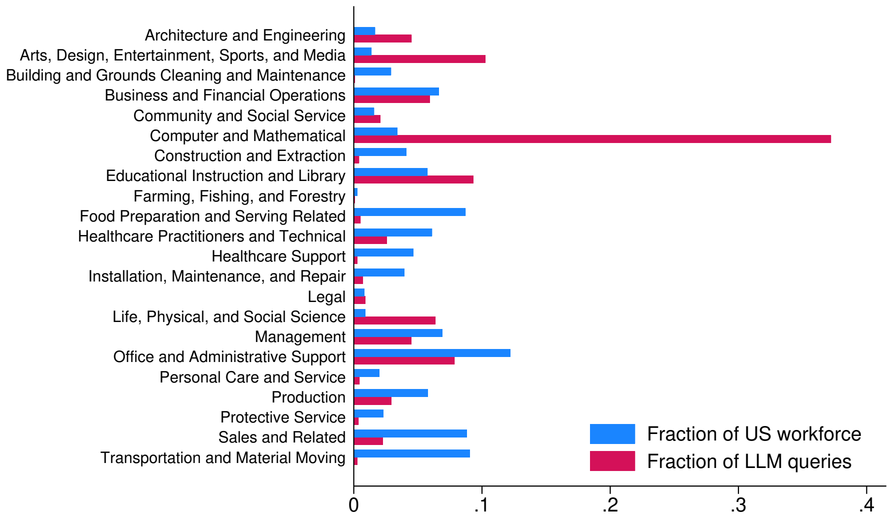
### 3.1 Anthropic Economic Index

Anthropic is one of the largest firms training LLMs and its Claude models are competitive with models offered by OpenAI and Google. Beginning in February 2025 Anthropic began releasing the "Anthropic Economic Index" which provides data on the composition of queries to their Claude models.[5] The data are organized by O*NET tasks; an instance of Claude is shown a user's query (in a privacy-preserving fashion) and asked which O*NET task it matches best. The resulting dataset shows which tasks are most commonly represented in interactions with Claude.[6]

Figure 1 shows, by broad occupational group, the share of Claude queries related to that occupation's tasks (red) and the share of that occupation in the workforce (blue). Strikingly, computer and mathematical occupations account for more that 1/3 of Claude queries, despite comprising only 3.4% of the workforce. Handa et al. (2025) show that these queries are essentially all computer programming-related. This simple fact motivates our focus on coding for the rest of the paper: Coders are clearly a very highly exposed group. This is also consistent with the survey data from Bonney et al. (2024) and Bick et al. (2024), which find the highest use rates among computer and mathematical occupations and in the information and professional and business services sectors.

[5]The analysis in this section is based on the first report from February, 2025. For the complete set of AEI releases, please see https://huggingface.co/datasets/Anthropic/EconomicIndex

[6]The data excludes API users, which might overrepresent coders, as well as business licenses: "*Because we focus on studying patterns in individual usage, the results shared in this paper excludes activity from business customers (i.e., Team, Enterprise, and all API customers).*"




*Note:* Fraction of US workforce and Claude queries associated with major SOC occupation groups
*Source:* Figure 3 of Handa et al. (2025)

Figure 1: Claude queries and workforce shares

Note that we do not have data on the composition of work-related queries going to other LLM tools, like ChatGPT or Gemini. Claude's consumer market share appears to be smaller that the major competitors, though the gap appears smaller for business users.[7] However, to a first order the capabilities of these models are similar across providers so we should expect broadly similar patterns to hold.[8] Massenkoff and McCrory (2026) show that with updated Anthropic data and an updated methodology computer programmers and software developers continue to be highly exposed. Independent of the AEI data, Bick et al. (2024) show that computer/mathematical occupations have the highest generative AI adoption rates, and Bonney et al. (2024) find that the information sector ranks highest in AI exposure.

[7]For consumer use see Figure A.1 here, for indicative evidence on business use see here.

[8]Chatterji et al. (2025) report that 4.2 percent of ChatGPT prompts across all types of consumer plans between 2024 and 2025 were related to computer programming, notably lower than what Handa et al. (2025) find in their Claude data. This may reflect higher use of ChatGPT for personal purposes. Handa et al. (2025) estimate less than a quarter of queries in their dataset are non-work.

## 3.2 What occupations are "coders"?

Having established that computer programming is a highly exposed skill, we need a principled way to define computer programming-intensive occupations. The computer and mathematical occupations major group is not ideal because it includes some occupations that are fairly removed from coding, such as "Information Technology Project Managers" and "Computer User Support Specialists." In addition, some coding-intensive jobs are found in other groups like "Computer Numerically Controlled Tool Programmers" and "Statistical Assistants."

We rely on data from O*NET. O*NET is a U.S. Department of Labor-sponsored database that provides standardized, detailed information on occupations, including required skills, tasks, knowledge, and work activities. O*NET data—in particular the tasks and work activities—have been widely used in labor economics, see Acemoglu and Autor (2011) and Eloundou et al. (2024) among many others. The tasks and work activities are very detailed and often specific to single occupations or a small grouping. Separately O*NET defines broader *skills* and rates their importance across most occupations. In particular O*NET has a computer programming skill rating, which measures the importance of the skill for the occupation on a 1-5 scale. This is a natural fit for our analysis. The skill variable is defined for about 900 of the roughly 1,000 O*NET occupations. For the remainder we impute the skill on basis of their tasks and related occupations using text embeddings and a random forest, see Appendix A for details.[9]

In the next section we show the most coding-intensive CPS occupations, but Table A3

[9]One special case is the occupation "Software developers." They are a large occupation—accounting for more than one percent of U.S. employment—and the occupation has been growing quickly. The importance of software developers is noted in this recent article. O*NET does not provide a computer programming skill value for software developers, and our imputation based on task descriptions assigns them a fairly low value. This is because the O*NET task list for software developers makes them seems like they only do a little coding: the tasks mentioned are things like designing software, meeting clients, etc.

Our sense is that modern software developers are very coding-intensive. Anecdotally the people that refer to themselves as "Software devs" or "software engineers" are hands-on coders and not, for example, planners that only tell others what to code. In Appendix A we provide evidence from self-reported job duties that this is the case. Consequently, we ignore the imputed skill variable for software developers and include them in the list of coding-intensive occupations.

| *CPS Code* | *Title* | *Employment Share* | *Cum. Emp. Share* | *Coding Importance* |
|---|---|---|---|---|
| 1010 | computer programmers | 0.288 | 0.288 | 4.750 |
| 1020 | software developers, applications and systems software | 1.369 | 1.657 | 3.986 |
| 1100 | network and computer systems administrators | 0.137 | 1.794 | 3.620 |
| 1060 | database administrators | 0.087 | 1.880 | 3.473 |
| 7900 | computer control programmers and operators | 0.049 | 1.930 | 3.120 |
| 5920 | statistical assistants | 0.011 | 1.940 | 3.000 |
| 1760 | physical scientists, nec | 0.279 | 2.219 | 2.880 |
| 1240 | mathematical science occupations, nec | 0.233 | 2.452 | 2.854 |
| 1700 | astronomers and physicists | 0.017 | 2.468 | 2.815 |
| 1000 | Comp. scientists and web dev. | 1.235 | 3.703 | 2.780 |
| 0110 | computer and information systems managers | 0.482 | 4.186 | 2.750 |
| 1220 | operations research analysts | 0.086 | 4.271 | 2.620 |
| 1460 | mechanical engineers | 0.255 | 4.526 | 2.543 |
| 1400 | computer hardware engineers | 0.057 | 4.583 | 2.500 |
| 1350 | chemical engineers | 0.047 | 4.630 | 2.380 |

*Note:* Top CPS occupations for coding importance sorted in descending order. Horizontal line shows our threshold for an occupation being coding-intensive. Employment share is the occupation's fraction of 2022 CPS employment in percentage points. Cum. Emp. Share is the cumulative employment share. Coding Importance is the O*NET computer programming skill metric, averaged within each CPS occupation when the CPS occupations are aggregations of O*NET occupations.
*Source:* CPS, O*NET, authors' calculations

Table 1: Coding-intensive CPS Occupations

shows coding intensity using the finer O*NET occupations.

### 3.3 Link to CPS

We link the O*NET data to CPS so we can define the programming skill variable on the universe of CPS occupations. CPS uses (a subset of) Census occupations codes, which can be crosswalked to Standard Occupational Classification (SOC) codes but are more coarse. O*NET uses occupation codes based on the SOC codes, though O*NET's codes are more detailed. More details of the linking process are found in Appendix B.

There are no CPS data for October 2025. We interpolate those industry-occupation employment counts from the September and November data.

Table 1 lists the most coding-intensive occupations in the CPS. The CPS code is `occ2010`, the longitudinally consistent occupation code provided by IPUMS. The third and fourth columns give each occupation's employment share and the cumulative employment share respectively. We see the top ranks are dominated by computer-oriented occupations. The Employment Share column makes clear the importance of software developers ("software developers, applications and system software"). Using a programming skill threshold of 2.76, we find that the coding-intensive occupations comprise about 3.7 percent of total employment. This threshold makes sure we exclude the explicitly management-oriented "computer and information systems managers" as well as mechanical engineers, both of which strike us as less coding-intensive than the higher-ranked occupations.

### 3.4 Industries

In addition to defining a coding-intensive group of occupations we define a group of industries where these occupations are over-represented.[10] Table 2 shows a sample of the CPS industries sorted by their intensity of coder employment: the share of their workforce that are coders ("Coders share of ind." in the table). The second column gives the industry's share of 2022 national employment, and the third column is the industry's share of coder employment. Finally, the last column shows the cumulative share of coders employed.

The top industry, "Computer systems design and related services" is more than 40 percent coders and accounts for more than 30 percent of national coder employment.[12] The other top industries are a mix of software/computer/high tech industries, with a mix of other sectors. We set the threshold for coder-intensive industries at 10 percent of industry employment. This covers nearly half of all coder employment in the U.S.

The industries in Table 2 are based on Census industry codes. Table A4 shows, for 2022, the NAICS codes that are mapped into each industry. The "NAICS percent of group" column is the fraction of the Census industry code employment that is accounted for by that

[10] IPUMS provides an "`ind1990`" longitudinally-consistent industry code based on the 1990 industry codes, which in turn are based on the 1987 SIC codes. These codes are not a great match given our focus on recent history and the tech sector.[11] Instead, we develop longitudinally-consistent codes linking together the 2012, 2017 and 2022 version of the Census industry codes. We link codes by coarsening each code that results in a many-to-one match, avoiding the need to allocate employment across split/merged codes.

[12] See Decker and Haltiwanger (2024) for more on entry dynamics in high tech and computer systems design in particular.

| Title | Coders share of ind. | Ind. share of emp. | Ind. share of coders | Cum. share of coders |
|---|---|---|---|---|
| Computer systems design and related services | 44.99 | 2.68 | 32.61 | 32.61 |
| Software publishers | 36.04 | .12 | 1.12 | 33.74 |
| Data processing, hosting, and related services | 32.78 | .1 | .85 | 34.59 |
| Scientific research and development services | 24.72 | .53 | 3.54 | 38.13 |
| Computer and peripheral equipment manufacturing | 21.32 | .07 | .4 | 38.53 |
| Other telecommunications services | 15.29 | .27 | 1.11 | 39.65 |
| Newspaper publishers | 11.74 | .57 | 1.81 | 41.45 |
| Pharmaceutical and medicine manufacturing | 11.41 | .43 | 1.33 | 42.78 |
| Electric and gas, and other combinations | 11.11 | .07 | .21 | 43 |
| Other general government and support | 10.19 | .1 | .26 | 43.26 |
| National security and international affairs | 10.17 | .66 | 1.8 | 45.06 |
| Aircraft and parts manufacturing | 10.07 | .54 | 1.46 | 46.52 |
| Electronic component and product manufacturing, n.e.c. | 9.75 | .41 | 1.08 | 47.6 |
| —-Misc.—-Professional, Scientific, and Management, and Administrative | 9.69 | .06 | .17 | 47.77 |
| Soap, cleaning compound, and cosmetics manufacturing | 9.51 | .09 | .24 | 48.01 |
| Banking and related activities | 9.35 | 1.39 | 3.51 | 51.52 |
| Communications, and audio and video equipment manufacturing | 9.33 | .06 | .16 | 51.68 |
| Non-depository credit and related activities | 9.22 | .77 | 1.93 | 53.61 |
| Navigational, measuring, electromedical, and control instruments manufacturing | 8.88 | .1 | .24 | 53.84 |

*Note:* Top CPS industries in terms of coder intensity, sorted in descending order. Employment share is the occupation's fraction of 2022 CPS employment in percentage points.
*Source:* CPS, O*NET, authors' calculations

Table 2: Industries sorted by intensity of coders

NAICS.

### 3.5 Other Measures of AI Exposure

Much of the recent literature on AI and the labor market uses the exposure measures calculated by Felten et al. (2023) and especially Eloundou et al. (2024). Eloundou et al. (2024) used GPT-4 to impute each occupation's generative AI exposure based on ONET detailed work activities. They confirmed that GPT-4's classifications largely align with crowdsourced human classifications.

Eloundou et al. (2024) is an impressive and influential early contribution. However, being developed so early it does have disadvantages. In particular, the exposure metrics predate the public release of GPT-4, 4o, all Claude models, subsequent "reasoning" models such as o3 and coding tools like Claude Code and Codex. Notably, the knowledge cutoff for GPT-4 was September 2021, well before even the release of ChatGPT and major image generation tools like Midjourney and Stable Diffusion. This means that GPT-4 had little in its training data to help it judge AI exposure, forcing it to rely heavily on the prompts where

Eloundou et al. (2024) describe generative AI capabilities.

We believe that real-world AI use data—like the AEI—has some important advantages over other exposure measures. By its nature, the AEI shows how people are actually using AI in the real world. At the same time, AEI also has some blind spots. It only covers *current* use and perhaps can't flag occupations which are likely to be highly exposed in the near future. It is also limited to the usage of the public version of the Claude chatbot, which could omit genAI exposure through, e.g., customized customer-service representative assistants and specialized image-generation tools. It could be that the Eloundou et al. (2024) metric is better suited to understanding which occupations are likely to eventually be exposed, while the AEI measure tells us which have already been exposed and thus may have measurable changes in outcomes.

To understand the measurement landscape we compare the "GPTs are GPTs" exposure measure (Eloundou et al., 2024)—which we call "GPTs exposure"—with the AEI exposure measure using ordinal rankings. For both metrics we divide occupations into (1) a most-exposed grouping accounting for about 20 percent of employment, and (2) the remainder.[13] This is motivated by the literature (e.g. Brynjolfsson et al. 2025) that focuses on the top quintile of AI exposure.

Table 3 shows the distribution of employment across these groups. Focusing on the first column we see that just over half of high-GPTs exposure employment is also classified as high-AEI exposure. Likewise, in the first row we see just over half the workers which are high-AEI exposure are also high-GPTs exposure. While it is encouraging that there is significant overlap between the GPTs and AEI metrics, nonetheless it is troubling that the two disagree about half the time as to which occupations are in the highest exposure grouping.

The first row of Table 4 shows what fraction of coding occupations are high-exposure under the AEI and GPTs groupings. The coding occupations are overwhelmingly in the high

[13] This is complicated by the fact that Handa et al. (2025) use ONET data based on 2010 SOC occupation codes and Eloundou et al. (2024) using 2018 SOC codes. We crosswalk the two to a common set of more coarse occupation codes taking the unweighted average of exposure within each coarse code.

| | (1) GPTs exposure high | (2) GPTs exposure low |
|---|---|---|
| AEI exposure high | 10.01 | 8.79 |
| AEI exposure low | 9.92 | 71.23 |

*Note:* Cells show the percent of U.S. private employment in each class. "GPTs exposure" uses Eloundou et al. (2024)'s GPT-$\beta$ metric, "AEI exposure" is based on Handa et al. (2025). High exposure occupations are those in the (approximate) employment-weighted top 20 percent of exposure, "low" exposure occupations are the remaining roughly 80 percent.
*Source:* Eloundou et al. (2024), Handa et al. (2025), OES, authors' calculations

Table 3: Distribution of Employment by generative AI exposure (Percent of total employment)

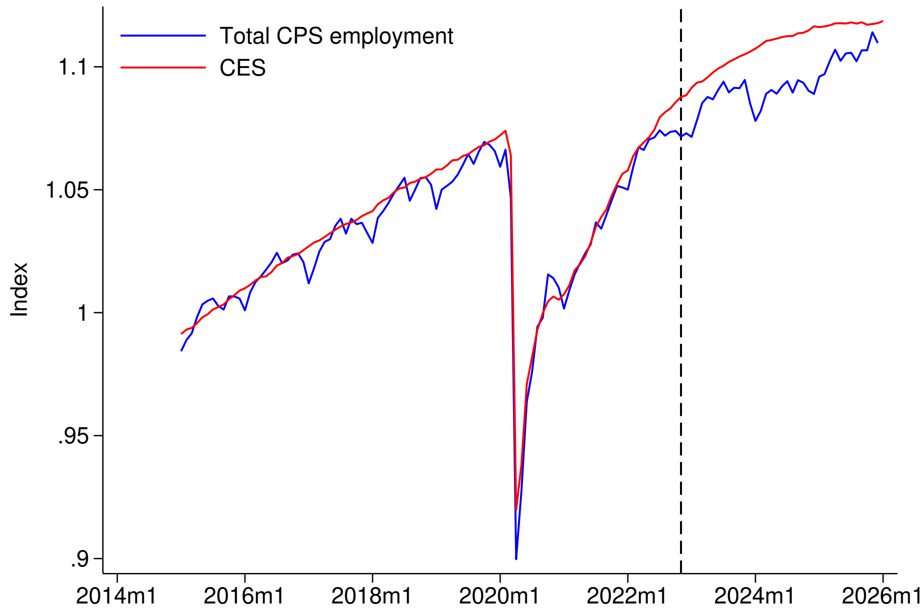
exposure group according to both metrics, with more than 98 percent of coder employment in the highest quintiles. Recall that the coding occupations grouping was motivated in part by examination of the AEI data, so it is not totally surprising that coding occupations tend to be highly exposed there. Regardless it is useful confirmation to see that the programming skill classification ends up agreeing both with the AEI metric and the (more independent) GPTs metric. For reference, the second row of the table shows that non-coding occupations are far less likely to be considered high exposure, with only about 17-18 percent of non-coders ending up in those groups.

In summary, we find—like Gimbel et al. (2025)—significant differences between Eloundou et al. (2024) and Handa et al. (2025) in terms of which occupations are highly exposed to generative AI though both metrics agree that our coder occupations are highly exposed.

| | (1) GPTs exposure high | (2) AEI exposure high |
|---|---|---|
| Coding occupations | 99.5 | 98.2 |
| Non-coding occupations | 18.0 | 16.9 |

*Note:* Row one shows the percent of coding employment that falls into the high exposure groups reported above. Row two shows the percent of non-coding employment falling into those groups. "GPTs exposure" uses Eloundou et al. (2024)'s GPT-$\beta$ metric, "AEI exposure" is based on Handa et al. (2025). High exposure occupations are those in the (approximate) employment-weighted top 20 percent of exposure, "low" exposure occupations are the remaining roughly 80 percent.
*Source:* Eloundou et al. (2024), Handa et al. (2025), OES, authors' calculations

Table 4: Percent of coding/non-coding employment falling into high-exposure groups




*Note:* Indexed levels of total CPS employment (NSA) and CES employment (SA). Vertical line indicates ChatGPT release date
*Source:* CPS, CES

Figure 2: CPS and CES indexed levels

### 3.6 Normalizations

When looking at employment levels it is important to remember that trend employment in CPS has sometimes been at odds with data from other sources. Figure 2 shows an index of (NSA) CPS employment with the published (SA) CES levels. Neither has been adjusted to account for the relatively small scope differences between the two. While CPS grew slightly slower than CES pre-covid, a larger gap has opened post-Covid with CPS only growing 0.3 percent between March 2023 and March 2024 while CES grew 1.2 percent during the same period. Note that CES is benchmarked to the QCEW comprehensive administrative data through March 2025 so it should not suffer from much error over this period (though we don't yet know how accurate it is for post-March 2025.)

To account for this divergence we calculate trends as employment *shares* in CPS and then convert them to levels using CES employment totals. This fairly crude adjustment ensures that our numbers are consistent with CES aggregate growth rates. In the absence of this correction many occupation groups would incorrectly show a deceleration in employment post-Covid.

# 4 Regression Specifications

Our basic approach is to test whether coder employment growth changed with the introduction of ChatGPT in November 2022. Our baseline specification is

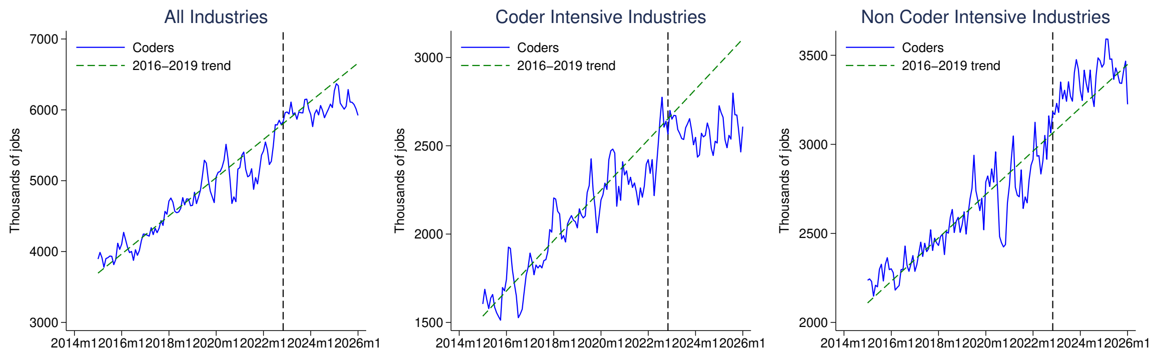
$$\ln e_t = \alpha + \beta_1 \cdot \mathbf{1}\{t \geq 2022m11\} + \beta_2 \cdot t + \beta_3 \cdot \mathbf{1}\{t \geq 2022m11\} \cdot t + \varepsilon_t \tag{1}$$

where $\ln e_t$ is log monthly employment of coders, $\beta_1$ shows the level shift in log monthly employment post-ChatGPT, $\beta_2$ is pre-ChatGPT trend growth, and $\beta_3$ captures the change in growth post-ChatGPT. We normalize the dependent variable so that the coefficients can be interpreted as annual growth effects in percentage points.

We run these regressions both for all industries and separately for coder intensive and non-coder intensive industries. In addition, we show robustness by dropping the Covid era (on the theory that labor markets were dislocated in this period) and alternatively dropping 2022 and 2023 (since any LLM/ChatGPT effect might not be dated precisely to November 2022.)

In Section 5.2 we develop a counterfactual employment series $Z_t$, meant to capture only shifts in employment due to industry-level shocks, not occupation-specific shocks. Then $\ln e_t - \ln Z_t$ is the occupation-specific shock to employment. To test whether ChatGPT generated an occupation-specific shock we run (1) with the new dependent variable:

$$\ln e_t - \ln Z_t = \alpha + \beta_1 \cdot \mathbf{1}\{t \geq 2022m11\} + \beta_2 \cdot t + \beta_3 \cdot \mathbf{1}\{t \geq 2022m11\} \cdot t + \varepsilon_t \tag{2}$$




*Note:* Employment level of coder occupations based on O*NET programming skill. Dashed line shows ChatGPT release date (November 2022.) Coder intensive industries are those where over 10 percent of workers are coders, non-coder intensive industries are the remainder.
*Source:* O*NET, CPS, authors' calculations

Figure 3: Coder employment

# 5 Results

Figure 3 shows coder employment in all industries, coder intensive industries, and non-coder intensive industries. Recall this is based on CPS employment shares, adjusted so that total CPS employment tracks CES employment. The vertical dashed line marks November 2022, the release date of ChatGPT. Our question is whether the employment of coders has been notably different after that date. Focusing on the first panel, there does appear to be a kink in employment growth, with pre-ChatGPT rapid growth flattening out. While employment is still fairly close to the pre-Covid linear trend the kink clearly suggests a change in employment dynamics around the introduction of ChatGPT. Turning to coder intensive industries in the middle panel, this pattern is more stark. Coder employment in these industries has been essentially flat since late 2022. The pattern in non-coder intensive industries is different with employment very close to the pre-Covid linear trend.

### 5.1 Initial Regressions

Table 5 shows the regression results. The first three columns show results for all industries, using the full sample, dropping the Covid era, and dropping 2022-2023 respectively.[14] The dependent variable is log employment times 1200, so the coefficients can be interpreted as annual growth rates. To account for autocorrelation we use Newey-West standard errors, allowing for 16 lags. This is long enough both to account for seasonality and rotation group issues (individuals leave the survey after 4 months in, an 8 month break, and a final 4 months in.) We see from the trend coefficient (the linear time trend) that coder employment averaged an annual growth rate of about 4.8 percent pre-ChatGPT. This is very fast; total private employment only grew at a pace of about 1.3 percent over this time. The coefficient of interest, "Post-GPT*Trend", is the change in trend growth post-ChatGPT, which shows a substantial slowing of growth.

Columns 2, 5, and 8 show that the results are not sensitive to dropping the Covid period. Further, columns 3, 6, and 9 show that the results survive dropping the years 2022 and 2023. This last exercise is relevant given the uncertainty about when exactly an AI shock could be said to arrive: Coding tools based on GPT 3.5 (like Github Copilot) were available well before ChatGPT, on the other hand it is likely that business reactions would have not been instantaneous. The exercise shows that the results are not sensitive to the exact timing of the shock.

The results for coder-intensive industries (columns 4-6) are even stronger than those for all industries. Pre-ChatGPT growth was about 6 percent annually, and became flat or negative post-ChatGPT (though again the No 2022/3 specification shows less impact.) Finally, the non-coder intensive sector shows somewhat smaller changes in growth. Taken together these regressions confirm the impression from Figure 3.

[14] The no Covid sample drops 2020m1 through 2021m6. Most adults were vaccinated by 2021m6 and masking/distancing guidance was being relaxed.

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | 4.882*** | 4.871*** | 4.814*** | 6.110*** | 5.981*** | 6.213*** | 3.883*** | 3.977*** | 3.680*** |
| | (0.270) | (0.304) | (0.412) | (0.464) | (0.489) | (0.631) | (0.207) | (0.204) | (0.288) |
| Post-GPT*Trend | -3.884*** | -3.872*** | -3.339*** | -6.248*** | -6.118*** | -4.063*** | -1.970*** | -2.064*** | -2.759** |
| | (0.371) | (0.410) | (0.871) | (0.680) | (0.750) | (0.805) | (0.549) | (0.593) | (1.105) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is monthly (log) coder employment in the relevant industries. Post-GPT period starts in November 2022.

Table 5: Regressions, no controls

## 5.2 Industry Shocks

The assumption implicit in Table 5 is that no other factors changed the trajectory of coder employment around 2022. This is a strong assumption, especially since coder employment is fairly concentrated in certain industries (see Section 3.4). Industry-level shocks are common, implying that occupational employment is partially a function of industry shocks. Indeed, the popular Bartik or shift-share instrument *depends on* industry-level shocks being fairly large in some sense (Borusyak et al., 2025; Goldsmith-Pinkham et al., 2020). It is possible that the industries coders work for happened to cut employment (growth) around 2022 for reasons unrelated to AI. If this were the case we should see coder employment track employment of the narrow industries they work in. Put differently, industries should not (on average) change the coder's *share* of industry employment, even if some industries shrink employment totals. This suggests a counterfactual: calculating what coder employment growth would have been on the basis of purely between-industry factors; if it was purely the exposure-weighted average of industry growth. Then the remainder—the within-industry component—is the occupation-specific shock.

Formally, we can start with the accounting identity

$$g_{\cdot,o,t} = \sum_{i=1}^{N} s_{i,o,t-1} g_{i,o,t} \tag{3}$$

where $s_{i,o,t-1} = \frac{e_{i,o,t-1}}{e_{\cdot,o,t-1}}$ is the fraction of occupation $o$'s employment that falls into industry $i$ in $t-1$, $g_{i,o,t}$ is the growth rate of employment in occupation $o$ and industry $i$, and $g_{\cdot,o,t}$ is the aggregate growth rate of occupation $o$. Adding and subtracting industry level growth we obtain

$$g_{\cdot,o,t} = \underbrace{\sum_{i=1}^{N} s_{i,o,t-1}(g_{i,o,t} - g_{i,\cdot,t})}_{\text{Within-industry}} + \underbrace{\sum_{i=1}^{N} s_{i,o,t-1}\, g_{i,\cdot,t}}_{\text{Between-industry}} \tag{4}$$

Here the within-industry component adds up the deviations of the occupation-industry growth rates from their corresponding industry rates. In other words, it counts the degree to which industries become more or less occupation $o$-intensive. The between industry component captures the exposure-weighted average of industry growth. This is what occupation $o$'s growth would have been if all industries kept constant occupation shares. This type of analysis has been used by, among many others, Acemoglu and Autor (2011) for industry/occupation margins in the context of polarization, Katz and Murphy (1992) for analyzing wages, Davis and Haltiwanger (1992) for decomposing excess job reallocation, and Jaimovich and Siu (2020) for accounting for job polarization.

Note that equation 4 is an identity—it holds by construction. Our identifying assumption is that in the absence of occupation-specific shocks the within-industry component is zero. This is true if (in the absence of occupation shocks) industries grow and shrink homothetically, keeping occupation shares constant. This appears to be a reasonable approximation, though a formal test is beyond the scope of this paper. In Section 7 we present evidence that in several historical episodes this decomposition has provided the correct intuition.

Industry-level shocks may take a number of forms: changes in demand for industry output, changes in industry TFP, changes to financial constraints, and regulatory or tax changes are all candidates. Any of these factors would (generally) scale industry labor demand up or down but would not affect the optimal occupational mix. Occupation-specific shocks, on the other hand, would be changes in technology or possibly regulations that change the optimal occupational mix.

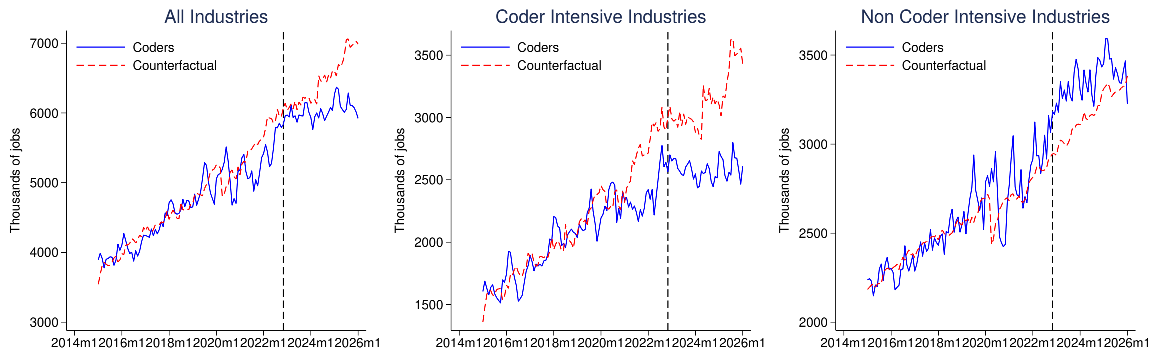
We cumulate the counterfactual growth rate $z_{\cdot,o,t} = \sum_{i=1}^{N} s_{i,o,t-1}\, g_{i,-o,t}$ into an index and normalize it to match the observed series in November 2022:

$$Z_{o,t} \propto \prod_{\tau=1}^{\tau=t} (1 + z_{\cdot,o,\tau}) \tag{5}$$

$$\propto \prod_{\tau=1}^{\tau=t} \left(1 + \sum_{i=1}^{N} s_{i,o,\tau-1}\, g_{i,-o,\tau}\right) \tag{6}$$

Note that this is not exactly the between-industry component of the accounting identity above: here we sum up the industry growth rate *excluding occupation o*, $g_{i,-o,t}$, rather than the actual industry growth rates $g_{i,\cdot,t}$. This conservative "leave one out" approach avoids the mechanical positive correlation between $z_{\cdot,o,t}$ and $g_{\cdot,o,t}$ when occupation shocks can influence industry growth.[15] Figure 4 shows the results. Focusing on the first panel, the counterfactual matches observed employment remarkably well up to the start of the Covid era. At that point the series diverge some, and—importantly—diverge further post-ChatGPT. This suggests that post-ChatGPT coder employment has not kept pace with employment growth in the relevant industries. A stronger version of this pattern can be seen in the second panel. Here the universe is restricted to industries where at least 10 percent of workers are coders ("coder intensive industries") and the exposure and counterfactual calculations are subject to the same restriction. The interpretation is that post-ChatGPT, even in this coder-intensive group of industries, industry employment grew faster than coder employment. The third panel shows that for non-coder intensive industries the counterfactual appears to follow observed employment better. However, note that the slope of the counterfactual is still greater than the observed slope post-ChatGPT.

[15] The decomposition is also dependent on the definition of the industries used. To avoid noise we require that an industry have at least 10 CPS respondents in each month. Industries failing this criterion are lumped into supersector-level "miscellaneous" groupings. The Appendix shows the results are not sensitive to the threshold.




*Note:* Employment level of coder occupations based on O*NET programming skill. Dashed line shows ChatGPT release date (November 2022.) Coder intensive industries are those where over 10 percent of workers are coders, non-coder intensive industries are the remainder.
*Source:* O*NET, CPS, authors' calculations

Figure 4: Coder employment with counterfactual

Table 6 shows the regression results for specification (2). The dependent variable is the difference between observed log employment and the counterfactual. This difference captures occupation-specific changes in employment: changes that cannot be explained by exposure to industry dynamics. Strikingly, the post-ChatGPT coefficients are all significant and clustered between -3 and -4 percent, with our baseline specification (column 1) showing an estimate of -3.23. The interpretation is that an occupation-specific shock has been suppressing coder employment growth by (conservatively) about 3 percent per year since the introduction of ChatGPT. This is consistent with ChatGPT leading to losses of (potential) coder jobs over that period.

One possibility is that other changes to the economy caused the growth rate of employment to change. After all, 2022 was still in the midst of post-Covid reopening. To address this possibility we run a placebo, picking a group unlikely to be affected by LLMs and seeing if their employment growth appears to change. We use the bottom (employment weighted) quintile of AI-exposed jobs from Handa et al. (2025) as this group. Table 7 shows that the post-GPT slope term is generally insignificant and has mixed signs across specifications. Only in the sample excluding 2022 and 2023 is it consistently negative and significant. Even

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -0.962*** | -1.184*** | -0.849* | -2.359*** | -2.797*** | -1.967*** | 0.616*** | 0.579*** | 0.533* |
| | (0.331) | (0.300) | (0.437) | (0.642) | (0.568) | (0.743) | (0.206) | (0.211) | (0.287) |
| Post-GPT*Trend | -3.222*** | -3.000*** | -4.427*** | -3.561*** | -3.123*** | -5.943*** | -3.016*** | -2.978*** | -3.410*** |
| | (0.635) | (0.598) | (0.850) | (1.072) | (0.998) | (1.162) | (0.414) | (0.420) | (0.818) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table 6: Regressions Controlling for Counterfactual

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -3.880*** | -3.782*** | -3.853*** | -2.997*** | -3.140*** | -2.614*** | -3.893*** | -3.779*** | -3.883*** |
| | (0.172) | (0.222) | (0.236) | (0.521) | (0.555) | (0.520) | (0.172) | (0.217) | (0.242) |
| Post-GPT*Trend | 0.328 | 0.229 | -0.955** | -1.435 | -1.292 | -2.897** | 0.409 | 0.296 | -0.864* |
| | (0.408) | (0.487) | (0.385) | (0.889) | (0.949) | (1.411) | (0.408) | (0.486) | (0.449) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table 7: Regressions Controlling for Counterfactual: Placebo group (low AI exposure)

in those specifications the coefficient is half the size or less than the corresponding coefficient from Table 6.

### 5.2.1 Discussion

The 3 percent effect size is large. We caution against interpreting this as a simple causal effect given the complexity of AI's potential economic effects and the measurement and identification challenges. However, to give some context we can do some back of the envelope calculations. Cumulating over the roughly 3 years since November 2022 and using 5.735 million coder jobs as the base value, the implication is that roughly 500,000 additional coder jobs would have existed in the absence of large-scale LLM use.[16,17] Brynjolfsson et al. (2025) do not provide an estimate of aggregate job losses due to AI, and their methodology is focused on measuring relative gaps between young and old workers. Nonetheless, treating their effect sizes as job losses we can approximate an aggregate effect. They find that among 22-25 year olds employment in the top two quartiles of AI exposure fell about 12 percent relative to employment in the bottom quartile. Starting from total private employment of 130 million, and assuming about 7.6 percent of the workforce is 22-25 years old (based on the CPS), a 12 percent job loss for two quartiles works out to about 475,000 jobs lost. In this sense a crude interpretation of our estimates are consistent with a crude translation of Brynjolfsson et al. (2025)'s estimates (which again are only estimates of relative job losses.)

For a number of reasons, we do not interpret the results as evidence that AI has eliminated 500,000 jobs from the economy. First, many coders would have found jobs—and possibly good, well-matched jobs—in other occupations. If occupational task mix and demand for other occupations is stable then those jobs may be not as good matches on average. But AI may be altering the task composition of occupations, such that a potential coder today may go into a management or other occupation that now uses more of their coding skills. In addition, these estimates—like those in other studies—do not attempt to account for the effects of automation on aggregate productivity and labor demand. In standard models the average worker is better off after a positive productivity shock, both from cheaper output and usually from increased aggregate labor demand, even if the displaced workers suffer persistent earnings losses. That noted occupational employment may fall, especially if AI is a substitute for labor or the demand for the relevant output is inelastic.

Even if coder employment is well below where it would have been, the estimates in the first three columns Table 5 show that coder employment has continued to grow post-ChatGPT. In this sense the raw statistics do not support a view that coder employment is

[16] Nonfarm employment was 158 million in November 2022, if the coder share is 3.7 that means there were about 5.8 million coders. A 3 percent annual decline for 3 years implies $5.8 \cdot 0.97^3 \approx 5.29$ million jobs remaining.

[17] Obviously, we are not likely see 500,000 additional unemployed coders in the data. Most of those workers would have been absorbed into other occupations.

collapsing. This is consistent with data from Indeed showing job openings for software developers mostly flat from 2024 and even edging up recently (after falling more than 50 percent over 2022-2023).

If firms have reduced coder employment relative to the counterfactual, it is not clear whether it is due to observed productivity gains or *anticipated* gains. In a model with fixed hiring and firing costs a firm that anticipates AI substituting for labor in the near future may freeze hiring and stay in the band of inaction for a time, even before actual productivity increases. Thus, declining employment growth does not imply actual productivity gains in the short run, though it strongly suggests they are at the minimum anticipated.

Examining the first panel of Figure 4, it can be seen that the gap between coder employment and the counterfactual does not open immediately after the release of ChatGPT. Indeed, the gap only widens significantly in the middle of 2024. This pattern appears plausible. It is unlikely that many firms would have instantly changed employment policy upon the announcement of ChatGPT; more likely it would have taken time and observation of model progress to commit to changes. However, the apparent timing is not quite consistent with a formal Bai-Perron test for structural breaks (Bai and Perron, 2003), as seen in Figure A1 The Bai-Perron test detects three breaks in the difference between monthly (log) coder employment and the counterfactual (log) employment series. This includes 2022m6 (just before the ChatGPT release), and 2020m3 (reflecting Covid). The test also finds a break at 2017m10, which might reflect sampling noise in the CPS and lacks a clear interpretation. While the break during Covid is expected, the break in 2022m6 is less intuitive. Taken at face value it could be evidence that firms were adjusting to the availability of Github Copilot (released October 2021) and the initial version of OpenAI Codex (released August 2021). Our sense is that this is unlikely, and it is more plausible that the Bai-Perron test is conflating an end-of-Covid break with a post-ChatGPT break. The noise in our relatively small samples make it hard to discern much more.

One question is to what extent AI is a *positive* demand shock for coders, and insofaras training and deploying LLMs require coders. We think this isn't likely to be a major factor yet. Taken together, employment at OpenAI, Anthropic, and Google Deepmind is likely under 15,000, and many of those workers are not coders.[18] Even if we multiply that by six to account for startups and the AI groups at Meta, Microsoft, and elsewhere we would still only account for less than two percent of U.S. coders.

#### 5.2.2 Tax law changes

The counterfactual controls for industry-level factors that may affect coder employment. However, a remaining threat to identification is occupation-specific factors unrelated to AI. One contemporaneous policy change was the implementation of the 2017 Tax Cut and Jobs Act (TCJA). Starting in 2022 this law meant that R&D expenses could no longer be expensed immediately and instead had to be amortized over several years. This change raised the effective cost of R&D spending, which is relevant because software development—including developer salaries—is typically considered an R&D activity. The Information sector accounts for about one quarter of R&D spending.[19] The TCJA implementation thus may have differentially affected coder hiring within the set of relevant industries.

Evidence on this channel is sparse. Cowx et al. (2025) find the law reduced R&D spending, while Du and Li (2025)—using a broader sample of firms and a longer time range—find no reduction in internal R&D and rising external R&D. Much of software R&D spending is done by the industries that are coder-intensive according to our classification. With this in mind it is encouraging that our results survive on the non-coder-intensive industry sample, where the TCJA changes would perhaps be less salient. Regardless, more research is likely needed.

[18] Public reports put OpenAI's employment around 4,000, Anthropic at 2,500, and DeepMind at 6,000.

[19] See NSF data here.

#### 5.2.3 Parametric assumptions

Ideally we would like to use a more local and/or non-parametric approach to identifying the effect of AI on jobs. Unfortunately, Covid likely renders this unworkable. The Covid-era labor market distortions mean that the period just prior to November 2022 is highly unusual, and unusual in ways we would expect to be transitory. Given these facts the next best option is to examine the pre-Covid period. Encouragingly for our approach, coder employment was growing more of less linearly pre-Covid; even better coder employment had reached its pre-Covid trend by late 2022. This is *suggestive* evidence that coder employment was on its historical trajectory and would have remained so in the absence of the LLM shock. While less than ideal, we think it better to look to the longer-run, pre-Covid trend for identification rather than rely on local trends around 2022 which will be distorted by Covid.

#### 5.2.4 Wages

If AI is decreasing demand for coders it may also be visible in wages paid. However, to the extent that the composition of workers changed (the main point in Brynjolfsson et al. (2025) and Lichtinger and Hosseini Maasoum (2025)) those wage changes may be swamped by having a larger share of the workers be older or more experienced. Looking that the log average wages of coders (winsorized at 1 percent) no break in 2022 is visible. While more work on this is likely necessary, we conclude that the main effect has been on the number of coders employed, not their wages.

#### 5.2.5 Robustness

The results in Table 6 show robustness across industry groups and sample period. However, there are other possibly relevant dimensions:

**Definition of "coders"** We adjust the threshold for coding-intensive jobs to make coders approximately 25 percent smaller and larger as a share of the workforce. In the stricter case this amounts to dropping a single large occupation from the group. Tables A5 and A6 show the results. Nearly all the signs remain negative, though statistical significance is lost in about the half the specifications. We take this as validation of the methodology. As described above we took considerable care to construct a grouping of occupations that is sufficiently large to allow for estimation and sufficiently focused to show any treatment effects, so we don't necessarily expect the results to survive changes to the definition of coders. To explore further we halve and double the size of the coder group in Tables A7 and A8. These tables show a similar pattern, with statistical significance often gone but most signs remaining.

**Software developers** As we noted, software developers are both a large occupational group and one where the O*NET task descriptions do not indicate very high coding intensity, though the survey respondent job duty write in information does indicate high coding intensity. Given the potential ambiguity of this group, Table A9 shows the results if we exclude software developers from the coders grouping. The results are qualitatively unchanged.

**Levels vs. shares** The main results are for CPS employment levels, scaled to match CES totals. Table A11 shows results for shares of CPS employment. In other words employment in each occupation (or occupation-industry pair) is first converted to a share of total CPS employment for the month, then the totals and counterfactuals are calculated. The results are exactly identical to the baseline, since the dependent variable is the log difference between observed employment and the counterfactual series. Conversion to shares means that employment and the counterfactual have both been divided through by a month-specific term, which cancels out when we log and take differences. Thus the dependent variable only differs by a constant.

**Growth calculations** Recall that equation 6 used a leave-one-out growth calculation, excluding coder employment when calculating industry level growth. In Table A10 we use total employment instead of leave-one-out and the results are qualitatively unchanged.

**Small industries** As noted in Section 5.2 the decomposition depends on the industry definitions. We experiment with the threshold below which industries get lumped into catch-all supersector groupings. The baseline threshold is to have at least 10 respondents in every month; we try setting the threshold to zero and 50 in Tables A12 and A13 show the results are qualitatively unchanged.

**E-commerce** The 2022 NAICS code revision eliminated the direct selling NAICS code and reclassified those establishments under the same NAICS as brick and mortar stores selling the same products. This change effectively switches e-commerce from its own NAICS industry to being pooled with a variety of physical retailers. Given that e-commerce is associated with the tech industry and coders it is possible this change would affect our results. Since the 2022 NAICS code change was implemented in CPS in 2025, we truncate the data in 2024 and reconstruct time-consistent industries through that year. Table A14 shows the results are qualitatively unchanged.

**Computer Systems Design** As noted above, 40 percent of coders work in computer systems design and related services (NAICS 5415). To check if the results are sensitive to the idiosyncracies of this industry we exclude it from the calculations. Table A15 shows the results are qualitatively unchanged.

**CPS weights** Each January the CPS updates population controls without reestimating the historical data, which creates a discontinuity. To see if this affects our results we use the smoothed weights developed by Coglianese et al. (2025). Table A16 shows the results are qualitatively unchanged.

## 6 Coder intensive industry dynamics

The previous section showed that industries substituted away from coders after the introduction of ChatGPT. The fact that narrow industries were substituting away from coders means that an industry-level shock cannot explain the coder employment declines. Nonetheless, it is informative to understand what industry-level influences have been in play recently and to what extent growth in the relevant industries is unusual.

Recalling that the coder intensive industries are dominated by software, software design, and internet-centric industries, there are a number of other factors that could explain declining employment growth in the sector. Interest rates began to rise in March 2022, potentially squeezing over-extended companies. Perhaps more importantly, it appears that the re-opening after Covid caused a reassessment of the sector, as spending from online services moved back to other forms of consumption. For example, advertising revenue fell at Google and Meta in 2022, consistent with the re-opening moving people away from online services. More tangentially, prior to 2022 there had been a wave of investor interest in cryptocurrencies but by 2022 the price of Bitcoin was falling, contributing to the collapse of FTX in November 2022 and generally increased pessimism about the prospects of blockchain technology. Finally, the treatment of R&D for tax purposes also changed in 2022—requiring domestic R&D spending to be amortized over five years rather than deducted immediately—though it is not clear if this change had any material impact.

In short, 2022 was a year that brought many changes to the coding intensive sector and estimating the contribution of each factor is beyond the scope of this paper. However, we can say something about level of employment relative to the pre-Covid trend.

For this exercise we use BLS's Occupation, Employment and Wage Statistics (OEWS), a survey of businesses that estimates the joint distribution of occupations across industries. Unlike CPS, it is low frequency: conducted semiannually and the published data averages 3 years worth of surveys. OEWS likely has more accurate estimates both because of a larger effective sample size (since it can target establishments that are larger) and because it relies

| *Industries* | *Pre-Covid* | *2020m2-2022m10* | *Post-ChatGPT* |
|---|---|---|---|
| Total private | 1.62 | .67 | .83 |
| Information (NAICS 51) | 1.21 | 2.41 | -2.76 |
| Coder-intensive agg. | 3.78 | 4.66 | -1 |
| Comp. Design (NAICS 5415) | 3.25 | 3.87 | -1.25 |
| Software (NAICS 5132) | 8.79 | 10.42 | .05 |

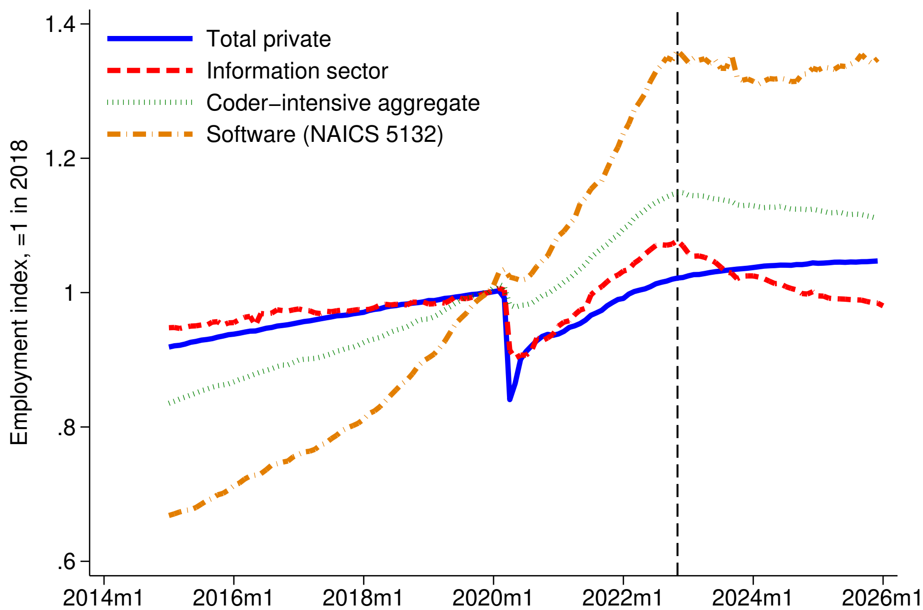
Table 8: CES industry growth

on the firm to report both occupation and industry (worker reporting of industry may be noisy). OEWS also reports data with somewhat more industry detail than CPS. In the OEWS we approximate our coder occupation group. Similar to our CPS methodology we define the coder intensive sector as those NAICS with a workforce of more than 10 percent coders.

With our coder intensive sector defined we link the data to the CES. CES reports monthly employment by detailed NAICS code. The sample is far larger than CPS making it a better gauge of employment levels.[20] CES also reports with time-consistent NAICS codes. Figure 5 shows indexed employment for a variety of industry groupings. It is evident that our coder intensive aggregate (green) has been been growing much faster than total private employment. The information sector is often used as a proxy for high-tech/computer-centric activity—this comparison shows the limits of that approach. Information (in red) was growing *slower* than total employment pre-Covid. That reversed during the Covid era, but declines since 2022 now put information below its 2019 level. Finally, for reference, the software industry (which is a component of the coding-intensive sector) was growing extraordinarily fast up until 2022. Table 8 shows growth of over 8 percent pre-Covid for the software industry followed by nearly flat growth post-ChatGPT.

We can explore this further in the CES data. Within CES we partition employment into 3 digit NAICS industries and a handful of 4 digit industries. Figure 6 shows a scatterplot of average post-2020 employment growth for these industries against average 2016-2019

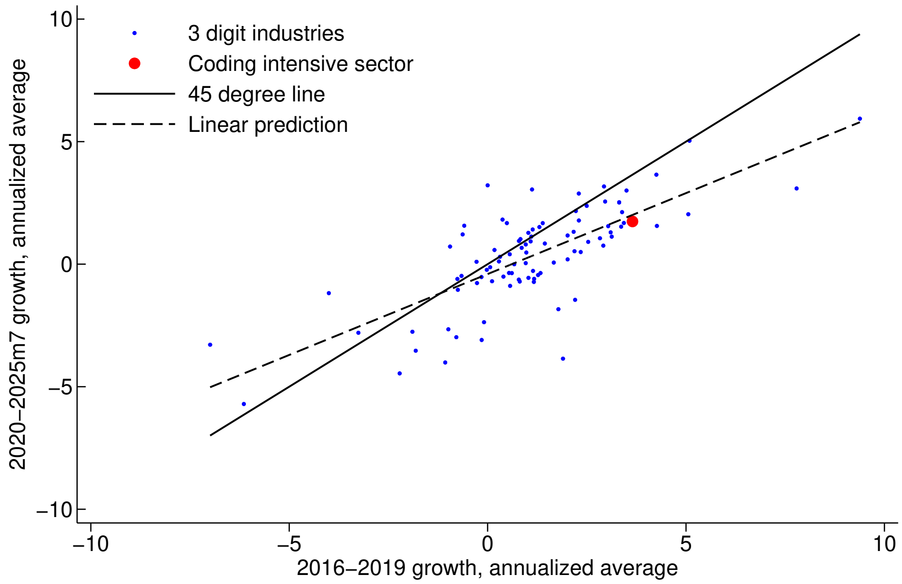
[20] CES is also benchmarked annually to the comprehensive QCEW data, meaning that except for most recent months the trends should be correct.




*Note:* Dashed line marks November 2022

*Source:* CES, OES, authors' calculations

Figure 5: Normalized CES Employment




*Note:* Annualized CES growth of 3 digit NAICS industries and the coding-intensive industry group.
*Source:* CES, OEWS, authors' calculations

Figure 6: Industry growth scatterplot

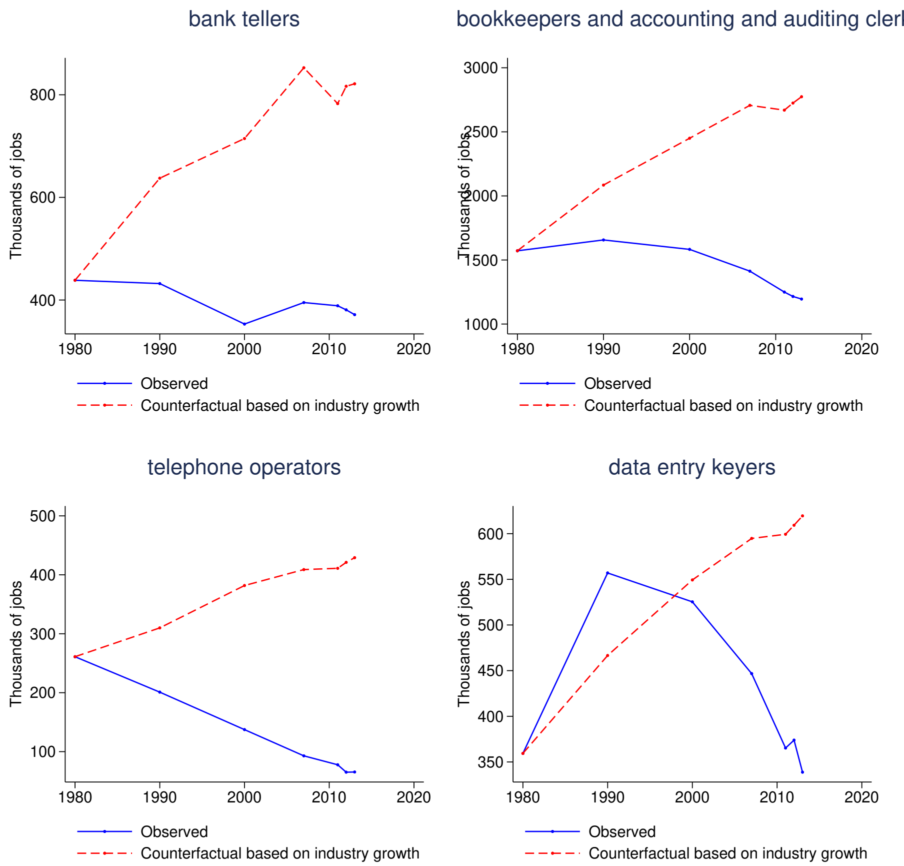
growth. The linear prediction line shows that there is some mean reversion: Industries that grew quickly pre-Covid continued to grow fast post-Covid, though not quite as fast. The red dot marks our coder intensive sector aggregate: it is exactly where we expect an industry to be given its pre-Covid trend. In other words coder intensive industry employment is right in line with where it "should" be, based on how industry employment evolved post-Covid. If the industry continues to have sluggish growth this will cease to be the case, but for the moment the level of employment is consistent with a return to trend following Covid overhiring.

# 7 Evaluating the Decomposition

Our results depend on the between-industry control variable correctly capturing industry-level shocks but not occupation-specific shocks. We can examine historical episodes to establish the reliability of the control. We use the data from Deming (2017), which is based on Census data from 1980 through 2012. The Census sample sizes are larger than the CPS. Importantly, the data have more or less consistent occupation and industry coding schemes over time.

For several example occupations we repeat the exercise from Section 5. The top left panel of Figure 7 shows the results for bank tellers. As discussed in Bessen (2015) and Autor (2015), the ATM substitutes for much of the work bank tellers originally did and one might expect the proliferation of ATMs to decrease bank teller employment. Instead, bank teller employment was flat from about 1980. This has been attributed to changes in regulation and the cost savings associated with ATMs, which made opening new branches (which still employ some tellers) more appealing. What is striking in the chart is the implication that bank tellers suffered a major negative occupation-specific shock. Basically, if bank tellers maintained their share of banking workforce they would have seen dramatic growth. This is completely consistent with ATMs (a technology that substitutes for labor) having a negative partial equilibrium impact on bank teller employment. What the decomposition cannot pick up is any general equilibrium relationship between occupation-specific shocks and industry growth. In summary, the graph correctly shows that there was a technological substitution away from bank tellers in the production function, though that same shock and other regulatory factors combined expanded the industry.

The second and third panels of Figure 7 show employment for bookkeepers and telephone operators. The literature on skill-biased technical change (e.g. Autor et al. (2003)) has argued that these types of routine cognitive occupations saw technological substitution in recent decades. This is borne out in the chart, where there is a large implied negative occupation-specific shock. The lower right panel shows the results for data entry keyers; this occupation initially grew quickly perhaps driven by computerization. But from the 1990s it suffered an occupation specific shock, possibly from the rise of the internet and the offshoring of these jobs. Importantly, the implication is that the industries that employed data

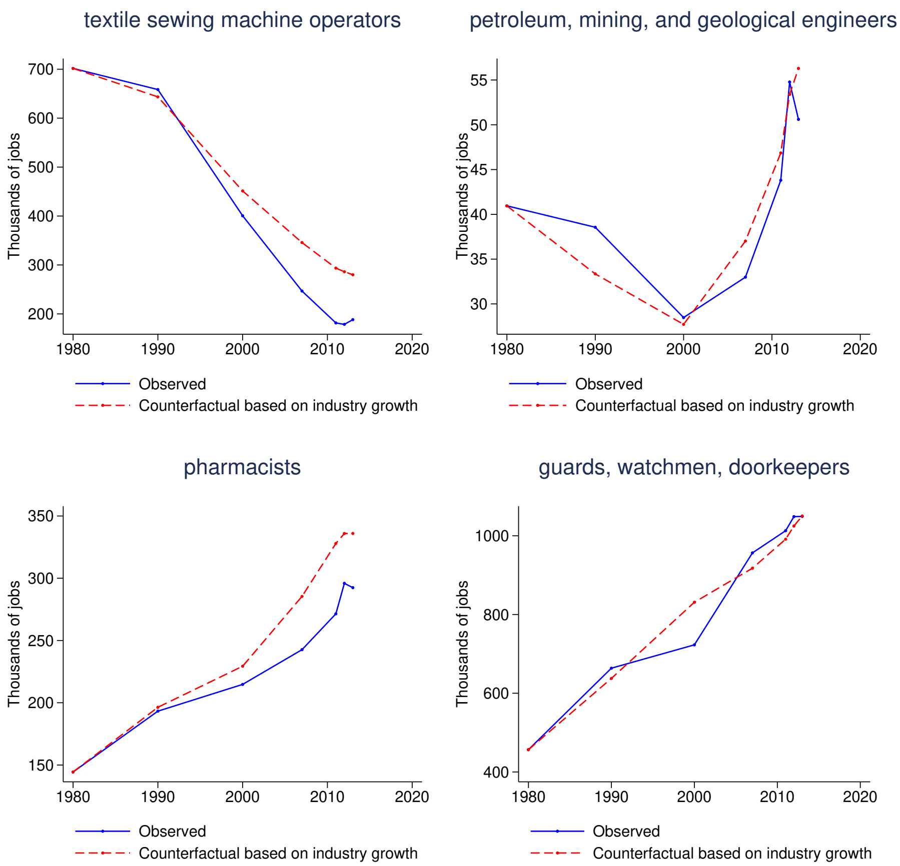



*Source:* ACS, Deming (2017), authors' calculations

Figure 7: Negative occupation-specific shocks

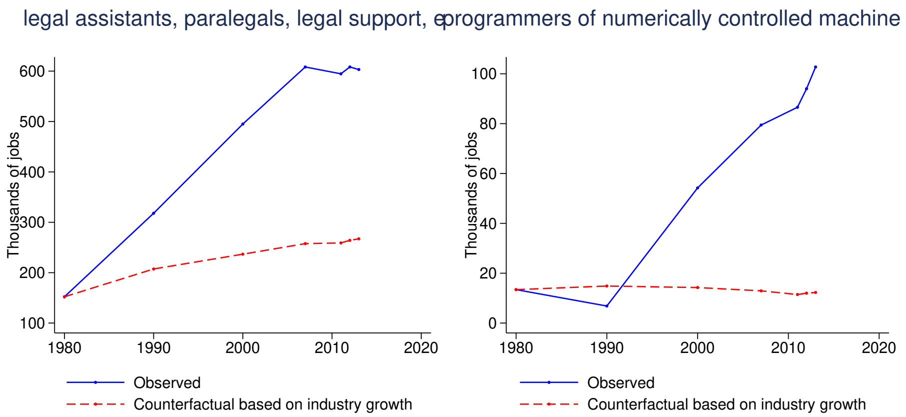
entry keyers mostly remained in the U.S. and only outsourced a limited set of occupations. Taken together, the examples in Figure 7 show that the between-industry decomposition can correctly identify occupation-specific shocks.

Figure 8 shows the analysis for examples of industry-level shocks. In the first panel we see sewing machine operator employment falling steadily (as trade eroded the domestic textile industry as a whole) and a smaller role for a negative occupation-specific shock. In the upper right panel we show petroleum, mining and geological engineers. This group saw employment first fall—as traditional oil production in the U.S. declined—and then reverse




*Source:* ACS, Deming (2017), authors' calculations

Figure 8: Examples of industry shocks




*Source:* ACS, Deming (2017), authors' calculations

Figure 9: Positive occupation-specific shocks

course as shale gas and shale oil became profitable (see Decker et al. (2016) and Decker et al. (2024) for more on shale oil). The counterfactual series attributes most of the employment dynamics to industry shocks, consistent with changes in oil demand and industry-level technology shocks that are neutral with respect to the types of labor employed. The bottom row of Figure 8 shows pharmacists and guards, two other occupations that have not experienced large occupation-specific shocks.

Finally, Figure 9 shows a couple examples of *positive* occupation-specific shocks. Since the 1980s legal assistants and paralegals have grown increasingly important in the legal domain. Similarly, computer numerically controlled (CNC) machine programmers have grown dramatically as the industry-based counterfactual would have predicted flat or declining employment. CNC machines are distinct from robots but their increased use is part of the continuing automation of manufacturing, especially in the motor vehicles and aerospace sectors.

These examples illustrate the utility of the between-industry decomposition, and demonstrate some cases where we can credibly identify other occupation-specific shocks.

## 8 Conclusion

The growth of LLMs has sparked considerable debate regarding the potential to automate complex tasks, potentially leading to lasting impacts on the labor market. This paper leverages occupation level data from O*NET and the CPS to measure the impact of ChatGPT's release in November 2022 on employment in computer programming, an occupation heavily exposed to AI. We find robust evidence that coder employment growth fell after that release. After controlling for industry-level shocks we find that coder employment growth has been 3 percent lower since the introduction of ChatGPT. This may reflect reallocation of tasks across occupations and we cannot control for all the relevant factors. Nonetheless, it suggests that AI is having a significant impact on this group of workers.

We believe it is important to control for industry-level shocks. Indeed, much of the debate over AI and the labor market has centered on industry-level stories such as interest-rate sensitivity and post-Covid dynamics. Our counterfactual exercise shows that the relevant industries differentially substituted away from coder employment. This is evidence that coders specifically experienced a shock not shared by their coworkers in other occupations. Our historical analysis shows many cases where the industry-based counterfactual correctly identifies the mix of industry and occupation-specific shocks affecting employment.

There are many unanswered questions. If firms are reducing coder employment, is it because they see concrete productivity gains, or are they holding back on hiring anticipating gains? Are remote or offshore coding services providers more vulnerable? In the longer run will the coder employment trend reverse, as new applications of (much cheaper) coding services develop? Are AI developments enabling the automation of other occupations (possibly with the input of coders)? To what extent are AI developments affecting aggregate productivity and labor demand? The current paper—suggesting a concrete employment effect on a well-specified group of workers—is only a first step toward answering these other questions.

## References


**Acemoglu, Daron**, "The simple macroeconomics of AI," *Economic Policy*, 2025, *40* (121), 13–58.

— **and David Autor**, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

— **and Jonas Loebbing**, "Automation and Polarization," *Journal of Political Economy*, March 2026, *134* (3). Electronically published January 22, 2026.

— **and Pascual Restrepo**, "Automation and new tasks: How technology displaces and reinstates labor," *Journal of economic perspectives*, 2019, *33* (2), 3–30.

**Ahn, Hie Joo and Nick Carollo**, "AI and Labor-Market Reallocation," 2026. Mimeo.

**Atkinson, Tyler and Shane Yamco**, "Young workers' employment drops in occupations with high AI exposure," January 2026.

**Autor, David H.**, "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, September 2015, *29* (3), 3–30.

— **, Frank Levy, and Richard J. Murnane**, "The Skill Content of Recent Technological Change: An Empirical Exploration*," *The Quarterly Journal of Economics*, 11 2003, *118* (4), 1279–1333.

**Bai, Jushan and Pierre Perron**, "Computation and analysis of multiple structural change models," *Journal of applied econometrics*, 2003, *18* (1), 1–22.

**Baily, Martin, David Byrne, Aidan Kane, and Paul Soto**, "Generative AI at the Crossroads: Light Bulb, Dynamo, or Microscope?," *arXiv preprint arXiv:2505.14588*, 2025.

**Bessen, James**, "Toil and Technology," *Finance & Development*, March 2015, *52* (1), 16–19.

**Bick, Alexander, Adam Blandin, and David J Deming**, "The Rapid Adoption of Generative AI," Working Paper 32966, National Bureau of Economic Research September 2024.

**Bonney, Kathryn, Cory Breaux, Cathy Buffington, Emin Dinlersoz, Lucia S Foster, Nathan Goldschlag, John C Haltiwanger, Zachary Kroff, and Keith Savage**, "Tracking firm use of AI in real time: A snapshot from the Business Trends and Outlook Survey," Technical Report, National Bureau of Economic Research 2024.

**Borusyak, Kirill, Peter Hull, and Xavier Jaravel**, "A practical guide to shift-share instruments," *Journal of Economic Perspectives*, 2025, *39* (1), 181–204.

**Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen**, "Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence," Technical Report 2025.

**_ , _ , and _** , "Canaries, Interest Rates, and Timing: More on the Recent Drivers of Employment Changes for Young Workers," February 2026. Online news/insight article.

**Chandar, Bharat**, "Tracking Employment Changes in AI-Exposed Jobs," June 2025. Working paper.

**Chatterji, Aaron, Thomas Cunningham, David J Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman**, "How People Use ChatGPT," Working Paper 34255, National Bureau of Economic Research September 2025.

**Coglianese, John M, Seth Murray, and Christopher J Nekarda**, "Harmonized Population and Labor Force Statistics," 2025.

**Cottier, Ben, Tamay Besiroglu, and David Owen**, "Challenges in Predicting AI Automation," 11 2023. Epoch AI blog.

**Cowx, Mary, Rebecca Lester, and Michelle L Nessa**, "The consequences of limiting the tax deductibility of R&D," 2025.

**Crane, Leland, Michael Green, and Paul Soto**, "Measuring AI Uptake in the Workplace," 2025.

**Davis, Steven J. and John Haltiwanger**, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

**Decker, Ryan A., Aaron Flaaen, and Maria D. Tito**, "Unraveling the Oil Conundrum: Productivity Improvements and Cost Declines in the U.S. Shale Oil Industry," FEDS Notes, Board of Governors of the Federal Reserve System March 2016.

**_ , Meagan McCollum, and Gregory B. Upton**, "Boom Town Business Dynamics," *Journal of Human Resources*, 2024, *59* (2), 627–651.

**Decker, Ryan and John Haltiwanger**, "High tech business entry in the pandemic era," 2024.

**Deming, David J.**, "The Growing Importance of Social Skills in the Labor Market*," *The Quarterly Journal of Economics*, 06 2017, *132* (4), 1593–1640.

**Dominski, Jacob and Yong Suk Lee**, "Advancing AI Capabilities and Evolving Labor Outcomes," 2025.

**Du, Wendi and Dongmei Li**, "Real Effects of External Innovation: Evidence from the Tax Cuts and Jobs Act," *Available at SSRN 5773322*, 2025.

**Eckhardt, Sarah and Nathan Goldschlag**, "AI and Jobs: The Final Word (Until the Next One)," August 2025. Online analysis/report.

**Eisfeldt, Andrea L, Gregor Schubert, and Miao Ben Zhang**, "Generative AI and firm values," Technical Report, National Bureau of Economic Research 2023.

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

**Felten, Edward W, Manav Raj, and Robert Seamans**, "Occupational heterogeneity in exposure to generative ai," *Available at SSRN 4414065*, 2023.

**Gimbel, Martha, Molly Kinder, Joshua Kendall, and Maddie Lee**, "Evaluating the impact of AI on the labor market: Current state of affairs," *The Budget Lab at Yale University*, 2025.

**Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift**, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

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

**Humlum, Anders**, "Robot Adoption and Labor Market Dynamics," November 2021. Manuscript.

__ **and Emilie Vestergaard**, "Large Language Models, Small Labor Market Effects," Technical Report, National Bureau of Economic Research 2025.

**Iscenko, Zanna and Fabien Curto Millet**, "The Trouble with Timing: Why the AI Displacement Narrative is Premature," Technical Report, Economic Innovation Group January 2026.

**Jaimovich, Nir and Henry E. Siu**, "Job Polarization and Jobless Recoveries," *The Review of Economics and Statistics*, 03 2020, *102* (1), 129–147.

**Katz, Lawrence F. and Kevin M. Murphy**, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

**Korinek, Anton and Donghyun Suh**, "Scenarios for the Transition to AGI," Technical Report, National Bureau of Economic Research 2024.

**Lichtinger, Guy and Seyed Mahdi Hosseini Maasoum**, "Generative AI as Seniority-Biased Technological Change: Evidence from US Resume and Job Posting Data," *Available at SSRN*, 2025.

**Massenkoff, Maxim and Peter McCrory**, "Labor Market Impacts of AI: A New Measure and Early Evidence," Report, Anthropic March 2026.

**Soto, Paul E**, "Research in Commotion: Measuring AI Research and Development through Conference Call Transcripts," 2025.

**Tomlinson, Kiran, Sonia Jaffe, Will Wang, Scott Counts, and Siddharth Suri**, "Working with AI: Measuring the Occupational Implications of Generative AI," 2025.

**Wiles, Emma and John J Horton**, "Generative ai and labor market matching efficiency," *Available at SSRN 5187344*, 2025.

# A Imputation of the Computer Programming Skill

For about 40 occupations we impute the skill variable on the basis of the occupation's tasks.[21] In particular, we concatenate the tasks into a string and get the embedding from the e5-large-v2 embedding model. Then we run a random forest to predict skill from the task embeddings using the sample of occupations with both task data and skill data. The rank correlation between actual and predicted skill is about 0.9 on this sample, indicating that the random forest does a good job translating tasks into programming skill requirements.

About 80 occupations are residual groups, like "machinists, all others." These occupations capture remainders and don't have task or skill data. We impute these using the related occupations based on the O*NET/SOC classification structure, where the occupations that are related share the first few digits in their O*NET/SOC code. For the remainder occupations we truncate the end of the occupation code until we find some occupations with data and then impute using the average.[22]

## A.1 Software Developers

The task-based skill imputation assigns software developers a computer programming skill importance of 2.28, less than that of business intelligence analysts and slightly above computer user support specialists. This is because the O*NET tasks for software developers only mention coding in passing and instead emphasize planning, requirements gathering, writing report, etc. Our sense is that actually many people that would call themselves software developers or software engineers (an occupation title that is mapped to software developers, see below) are doing work O*NET would associate with computer programmers: writing code and attending meetings to design and troubleshoot code. In the CPS data occupations

[21] Skills and importance ratings are based on surveys of incumbent workers in the occupation as well as assessments by occupational experts and analysts. For further information on the skills rating procedure, please see https://www.onetcenter.org/dl_files/AOSkills_Proc.pdf

[22] The exception is occupation 43-9199, "all other office workers." If we impute using truncated codes it matches them with "statistical assistants", a highly specialized and technical occupation. In this case we truncate further so the remainder matches with a broader set of office workers.

are supposed to be assigned based on the tasks done, not what the worker calls themselves, but we suspect measurement error could play a role here.

In Census-fielded surveys like the CPS respondents are asked for their job title and their usual duties. On the basis of this information analysts assign the respondent's occupation using the Census occupation codes (which can be crosswalked to SOC codes and O*NET occupations). Remarkably, Census provides a public-use microdata sample of written-in job titles and duties, as well as the Census occupation that was assigned.[23] These data are from the 2018 American Communities Survey (ACS), but should be comparable to the CPS data. Table A1 lists the written-in job duties for people coded as software developers (upper panel) and computer programmers (lower panel). There are many more examples for software developers because it is a more common occupation label. Software developers frequently mention writing code, which suggests they may be fairly similar to computer programmers. On the other hand, they more frequently use terms like "develop software" which (if the O*NET description is correct) could mean things other than coding.

To examine this in more depth we train a random forest on embeddings of the self-reported job duties to classify the duty (again using e5-large-v2 for embeddings). We manually label the duties for computer programmers, software developers, and a sample of other occupations. We label three classes: computer programming, software development, and non-programming/non-development. For this exercise we intentionally treat "software development" as an ambiguous class distinct from coding. Nearly every write-in for computer programming (from Table A1) is hand labeled as programming, while only about 1/3 of the software development observations are.

The random forest has a good fit on the hand labeled data and in the complete dataset the write-ins it classifies as programming are sensible. For each write-in observed we take the most likely class as the label and calculate the fraction of an occupation's write-ins that are classified as computer programming. Table A2 shows the occupations that are most

[23]See here.

programming-intensive according to this metric. Software developer is near the top of the ranking, only behind a few small occupations with very few observations and noisy estimates. This shows that software developers are among the closest occupations to computer programming in terms of intensity of computer programming, *even if we don't consider "software development" and "developing software" to include coding*. This makes us more confident that software developer duties are very close to those of programmers. With this in mind we override the task-based imputation and hard-code a 4 for the computer programming skill variable for software developers.

| |
|---|
| **Software Developer**<br>WORKED ON RESPONSIVE DESIGN AND DEVELOPED A SINGLE RESPONSIVE WEBSITE, U S GOVERNMENT, DESIGN DATACENTERS FOR STATE LOCAL GOVERNMENT AND HIGHER EDUCATION INSTITUTIONS, C4ISR SYSTEM DESIGN, COMPUTER WORK, DESIGN AND DEVELOP SOFTWARE FOR DEFENSE SYSTEMS, SYSTEMS ENGINEERING, VALIDATING PROGRAM REQUIREMENTS, SYSTEM ARCHITECTURE, TECHNOLOGY, DEVELOP METRICS AND ANALYTICS SOLUTIONS E G DASHBOARDS REPORTS ETC, DEMO PRODUCTS ASSIST TROUBLESHOOTING FOR PRODUCT TRIALS PROVIDE CUSTOMIZATIONS FOR THE CUSTOMER, TECHNICAL SUPPORT FOR HR SYSTEMS AND INTERFACES, AUTOMATED REGRESSION TESTING FOR SOFTWARE DEVELOPMENT, BUILD EDUCATIONAL WEB SOFTWARE, COMPUTER PROGRAMMER, CONFIGURE AND SUPPORT CLIENT SOFTWARE ENVIRONMENTS, CREATE WEBSITES AND APPS ONLINE, DESIGN AND DEVELOP SOFTWARE, DESIGN AND WRITE COMPUTER SOFTWARE, DESIGN ARCHITECT WRITE SOFTWARE AND LEAD TEAM, DESIGN DEVELOP AND TEST SOFTWARE APPLICATIONS, DESIGN SOFTWARE ARCHITECTURE AND IMPLEMENT SERVICES, DESIGNING CODING TESTING MAINTAINING COMPUTER SOFTWARE, DEVELOP AND MAINTAIN TECHNOLOGY MONITORING SOFTWARE, DEVELOP COMPUTER PROGRAMS TO MANAGE WEB SERVICE PLATFORM, DEVELOP SOFTWARE, DEVELOP SOFTWARE, DEVELOPING SOFTWARE, LEAD GRAPHICS ENGINEER, MAINTAIN GOVERNMENT SATELLITE PROGRAMS, MAINTAIN SYSTEMS AND SOFTWARE APPLICATIONS ASSOCIATED WITH VEHICLE RENTAL, PROGRAMMING, SOFTWARE ENGINEERING, SOFTWARE ENGINEERING, SOFTWARE PROGRAMMING IN PL/SQL JAVA ETC, SOFTWARE SERVICE, TESTING SOFTWARE, WORK ON COMPANY WEBSITE, WRITE COMPUTER CODE, WRITING CODE DESIGNING SOFTWARE SYSTEMS REVIEWING CODE, WRITING SOFTWARE, ENGINEER SOFTWARE, FOLLOW BEHAVIOR DRIVEN DEVELOPMENT METHODOLOGY TO DEVELOP SOFTWARE USE JAVA J BEHAVE CD/CI CONCEPTS, DEVELOPE SOFTWARE, PROGRAMMING, SOFTWARE DEVELOPMENT, CREATE AND MAINTAIN SOFTWARE USER INTERFACES REPORTS AUTOMATED PROCESSES, DEVELOP SOFTWARE, DEVELOP SOFTWARE, DEVELOP SOFTWARE FOR CLIENTS, EVALUATE BUSINESS NEEDS THEN DESIGN AND DEVELOP SOFTWARE SOLUTIONS, SOFTWARE DEVELOPMENT, SOFTWARE DEVELOPMENT AND TRAIN NEW HIRES, VERIFY TRADING APPLICATION, WEB DEVELOPMENT, WRITE CODE, WRITE CODE, WRITE CODE FOR MOBILE APPS, WRITE TEST AND MAINTAIN SOFTWARE, DEVELOP AND SUPPORT MARKETING CAMPAIGNS AND REPORTING SOFTWARE, HELP TO BUILD AND INSTALL SOFTWARE PRODUCT FOR CLIENT, DOCUMENT SOFTWARE REQUIREMENTS FOR CLIENTS AND VALIDATE DELIVERED PRODUCTS, DESIGN AND DEVELOP SOFTWARE DATABASE SYSTEMS, DEVELOP SOFTWARE, IMPLEMENTATION OF SOFTWARE TO PROVIDE ELECTRONIC CREDENTIALS TO LEARNERS AND SCHOOLS, CREATE SOFTWARE, SOFTWARE ON SATELLITE, TESTING SOFTWARE, PROGRAMMING ENHANCEMENTS, WRITE COMPUTER CODE & MONITOR COMPUTER SERVICES, TRACK COMPANY PROGRESS FOR PROJECTS, HELP CUSTOMERS AGENCIES WITH THEIR PAYROLL QUESTIONS AND PROJECTS, WRITE REQUIRMENT DOCUMENTS THAT TELL COMPUTER PROGRAMMERS WHAT TO PROGRAM, IT STRATEGY AND SOFTWARE IMPLEMENTATION, MAINTAINING SYSTEMS IN THE CASINO RESORT, INFRASTRUCTURE DESIGN IMPLEMENT AND MAINTAIN, SPEC PROVISION MAINTAIN UNIX AND LINUX SERVERS, WRITE WEB APPLICATIONS, SALES, ENGINEER IT SYSTEMS, DEVELOPING IAM TOOLS, DEVELOP CODE THAT ALLOWS ROBOTS TO MOVE, QUALIFICATION OF PRODUCTS, CODE SOFTWARE, DEVELOP SOFTWARE FOR NETWORKING DEVICES, WRITE SOFTWARE, CODING, WRITE FULL STACK CODE AND PLAN SOFTWARE STRATEGY FOR SERVICES-, DEVELOP SOFTWARES, CODING, DESIGN AND ANALYZE COMPUTER PROGRAMS, ANALYZE PROBLEMS AND CREATE COMPUTER PROGRAMS, COMPUTER PROGRAMMING, PUBLIC AND CONGRESSIONAL AFAIRS, DEVELOP CLINICAL APPLICATIONS FOR MASS SPECTROMETRY PRODUCTS AND MARKET TO APPROPRIATE CLIENTS, PLAN AND BUILD SOFTWARE, CODE DEVELOPMENT FOR INFORMATION SYSTEMS, INFORMATION DEVELOPING, SUPPORT AND ENHANCE SOFTWARE DEVELOPMENT, ANALYSIS |
| **Computer Programmer**<br>SAME, PROGRAMMER, MANAGE STUDENTS TUTORS AND WRITE CURRICULUM, DESIGN AND IMPLEMENT SOFTWARE, DEVELOPING SOFTWARE/ONLINE FOR COMPANIES, PROGRAM, SUPPORT, IT, PROGRAMMER, PROGRAMMING, SOFTWARE PROGRAMMER, COMPUTER PROGRAMMER, AUTOMATION PROGRAMMING, COMPUTER PROGRAMMER, CODING AND BUSINESS ANALYTICS, COMPUTER PROGRAMMING, COMPUTER PROGRAMMING, CREATE PROGRAMS AND REPORTS, DESIGN AND CONSTRUCT COMPUTER PROGRAMS, PROGRAMMING, WRITE AND MAINTAIN CLIENT SOFTWARE SYSTEM, DOING CODING FOR A CONSTRUCTION COMPANY |

Table A1: Self-reported job duties

| *Code* | *Title* | *Fraction Programming* | *Number of Obs.* |
|---|---|---|---|
| 1010 | Computer programmers | 0.727 | 22 |
| 5920 | Statistical assistants | 0.500 | 2 |
| 1240 | Other mathematical science occupations | 0.364 | 11 |
| 1031 | Web developers | 0.333 | 6 |
| 1021 | Software developers | 0.277 | 101 |
| 1800 | Economists | 0.167 | 6 |
| 7905 | Computer numerically controlled tool operators and programmers | 0.091 | 11 |
| 1220 | Operations research analysts | 0.077 | 13 |

*Note:* Top Census occupations by fraction of ACS duty write-ins classified as "computer programming". "Fraction Programming" givens the fraction of that occupation's duty write-ins classified as programming by a random forest run on the text embeddings. Number of Obs. is the number of respondents in that occupation with write-in data in the public-use file.
*Source:* ACS, authors' calculations

Table A2: Fractions of duty write-ins classified as computer programming by occupation

Table A3 shows the most coding-intensive occupations in O*NET. The "Importance" variable is the programming skill variable we use, which is designed to capture how critical having the skill is for the occupation. "Level" is a separate measure that captures the degree of computer programming sophistication needed for the occupation, it comoves very closely with Importance. O*NET explains the difference using the the example of lawyers and paralegals with respect to the speaking skill. Speaking is an important skill for both occupations, but level required is higher for lawyers, who often need to argue cases in formal settings. The ordering appears sensible, with the top ranks dominated by jobs that are obviously coding-intensive. Lower down in the ranks we start seeing slightly less computer-focused occupations, e.g. Physicists and Biostatisticians.

We choose a threshold value of 2.76 to define our programming-intensive, or *coder*, occupation group. This is low enough to include computer scientists and web developers (clearly coding-heavy jobs) but excludes, e.g., automotive engineers and information system managers.

| *Code* | *Title* | *Importance* | *Level* | *Imputation Type* |
|---:|---:|---:|---:|---:|
| Computer Programmers | 15-1251.00 | 4.750 | 4.880 | - |
| Web Developers | 15-1254.00 | 4.120 | 4.250 | - |
| Computer Numerically Controlled Tool Programmers | 51-9162.00 | 4.120 | 4.120 | - |
| Video Game Designers | 15-1255.01 | 4.000 | 3.880 | - |
| Software Developers | 15-1252.00 | 4.000 | 4.000 | C |
| Computer Network Architects | 15-1241.00 | 3.880 | 3.620 | - |
| Data Warehousing Specialists | 15-1243.01 | 3.750 | 3.750 | - |
| Software Quality Assurance Analysts and Testers | 15-1253.00 | 3.620 | 3.750 | - |
| Network and Computer Systems Administrators | 15-1244.00 | 3.620 | 3.880 | - |
| Computer and Information Research Scientists | 15-1221.00 | 3.620 | 4.500 | - |
| Computer Systems Engineers/Architects | 15-1299.08 | 3.500 | 4.250 | - |
| Biostatisticians | 15-2041.01 | 3.380 | 4.000 | - |
| Database Architects | 15-1243.00 | 3.380 | 4.120 | - |
| Physicists | 19-2012.00 | 3.380 | 3.880 | - |
| Database Administrators | 15-1242.00 | 3.380 | 3.880 | - |
| Clinical Data Managers | 15-2051.02 | 3.250 | 3.620 | - |
| Computer Systems Analysts | 15-1211.00 | 3.250 | 4.000 | - |
| Bioengineers and Biomedical Engineers | 17-2031.00 | 3.120 | 3.380 | - |
| Robotics Engineers | 17-2199.08 | 3.120 | 3.880 | - |
| Bioinformatics Technicians | 15-2099.01 | 3.120 | 3.000 | - |
| Web Administrators | 15-1299.01 | 3.120 | 2.880 | - |
| Mathematical Science Occupations, All Other | 15-2099.00 | 3.120 | 3.000 | B |
| Statisticians | 15-2041.00 | 3.000 | 3.120 | - |
| Computer Science Teachers, Postsecondary | 25-1021.00 | 3.000 | 3.120 | - |
| Statistical Assistants | 43-9111.00 | 3.000 | 3.500 | - |
| Physical Scientists, All Other | 19-2099.00 | 2.880 | 2.750 | B |
| Remote Sensing Scientists and Technologists | 19-2099.01 | 2.880 | 2.750 | - |
| Robotics Technicians | 17-3024.01 | 2.880 | 2.750 | - |
| Geographic Information Systems Technologists and Technicians | 15-1299.02 | 2.880 | 3.120 | - |
| Health Informatics Specialists | 15-1211.01 | 2.880 | 3.120 | - |
| Industrial-Organizational Psychologists | 19-3032.00 | 2.750 | 2.880 | - |
| Automotive Engineers | 17-2141.02 | 2.750 | 2.880 | - |
| Search Marketing Strategists | 13-1161.01 | 2.750 | 2.120 | - |
| Biologists | 19-1029.04 | 2.750 | 3.620 | - |
| Computer and Information Systems Managers | 11-3021.00 | 2.750 | 2.620 | - |
| Electronics Engineers, Except Computer | 17-2072.00 | 2.750 | 2.880 | - |
| Remote Sensing Technicians | 19-4099.03 | 2.750 | 3.250 | - |
| Logistics Engineers | 13-1081.01 | 2.620 | 2.250 | - |
| Desktop Publishers | 43-9031.00 | 2.620 | 2.620 | - |

*Note:* Top O*NET occupations for coding importance sorted in descending order. Horizontal line shows our threshold for an occupation being coding-intensive. Imputation type "A" means random forest-based using task embeddings (not shown), imputation type "B" means imputation using the truncated SOC code occupational groupings, imputation type "C" means the software developers hardcode based on written in job duties.
*Source:* CPS, O*NET, authors' calculations

Table A3: Coding-intensive ONET Occupations

## B Link to CPS

BLS provides crosswalks between O*NET and the SOC codes used in the "National Employment Matrix" (NEM), the industry-occupation employment totals they publish from the OES. The NEM codes are a modification of the SOC codes. Crosswalking from O*NET to NEM only results in the loss of 19 O*NET occupations, mostly military-related. When multiple O*NET occupations match to a single NEM occupation we take the simple average of the skill variable. We end up with 832 usable NEM occupations.

BLS also has a crosswalk between NEM codes and CPS occupation codes. Again we average the skill variable when multiple NEM occupations match one CPS occupation. We end up with 525 useable CPS occupations.

The O*NET and CPS codes we match to are based on the 2018 SOC, which the CPS data use from 2020. To get consistency before 2020 we need to use the standardized 2010 occupation codes that IPUMS has developed, `occ2010`. All workers in all CPS data post-2010 have `occ2010` codes on a consistent basis.

In the post-2020 sample, employed individuals have both `occ2010` and 2018-vintage occupation codes. We average to get a weighted crosswalk between the 2018 vintage and `occ2010`, which lets us assign programming skill to `occ2010` codes as the employment-weighted average of the occupations matching to them post-2020.

There are a small number of `occ2010` codes that were in use prior to 2020 but not afterwards. These codes don't appear in the weighted crosswalk above but they account for a (small) portion of pre-2020 employment. To get a skill variable for about 50 occupations fitting this description we use the official 2010-2018 Census occupation code bridge, and see if we can match the `occ2010` values to the NEM codes. After this step there are still a handful of occupations we can't populate, we drop these holdouts.

# C Industries

| *Census Industry* | *NAICS title* | *NAICS codes* | *NAICS percent of group* |
|---|---|---|---|
| Computer systems design and related services | Computer systems design and related services | 5415 | 100 |
| Software publishers | Software publishers | 5112 | 100 |
| Data processing, hosting, and related services | Data processing, hosting, and related services | 5182 | 100 |
| Scientific research and development services | Scientific research and development services | 5417 | 100 |
| Computer and peripheral equipment manufacturing | Computer and peripheral equipment manufacturing | 3341 | 100 |
| Other telecommunications services | Telecommunications, except wired telecommunications carriers | 517 exc. 517311 | 100 |
| Newspaper publishers | Newspaper publishers | 51111 | 13.79 |
| Newspaper publishers | Periodical, book, and directory publishers | 5111 exc. 51111 | 19.34 |
| Newspaper publishers | Broadcasting (except internet) | 515 | 34.4 |
| Newspaper publishers | Internet publishing and broadcasting and web search portals | 51913 | 28.47 |
| Newspaper publishers | Other information services, except libraries and archives, and internet p... | 5191 exc. 51912, 51913 | 4.01 |
| Pharmaceutical and medicine manufacturing | Pharmaceutical and medicine manufacturing | 3254 | 100 |
| Electric and gas, and other combinations | Electric and gas, and other combinations | Pts. 2211, 2212 | 100 |
| Other general government and support | Other general government and support | 92119 | 100 |
| National security and international affairs | National security and international affairs | 928 | 100 |
| Aircraft and parts manufacturing | Aircraft and parts manufacturing | 336411, 336412, 336413 | 100 |

Table A4: Industries sorted by intensity of coders, with NAICS codes

# D Robustness

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | 0.0359 | -0.233 | -0.0529 | -1.847*** | -2.254*** | -1.578*** | 1.777*** | 1.632** | 1.455** |
| | (0.356) | (0.470) | (0.478) | (0.376) | (0.307) | (0.448) | (0.495) | (0.658) | (0.580) |
| Post-GPT*Trend | -2.827*** | -2.557*** | -1.253* | -1.055 | -0.648 | -3.246** | -4.057*** | -3.911*** | 0.0741 |
| | (0.503) | (0.607) | (0.747) | (0.903) | (0.867) | (1.438) | (1.054) | (1.218) | (0.881) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A5: Regressions Controlling for Counterfactual, 25 percent stricter definition of "coder"

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -1.800*** | -1.972*** | -1.661*** | -2.722*** | -2.870*** | -2.779*** | 0.317 | 0.0968 | 0.850* |
| | (0.151) | (0.129) | (0.200) | (0.230) | (0.229) | (0.324) | (0.390) | (0.457) | (0.454) |
| Post-GPT*Trend | -1.897 | -1.725 | -5.453*** | -3.017** | -2.869** | -7.463*** | 0.121 | 0.341 | -1.709 |
| | (1.148) | (1.179) | (1.446) | (1.305) | (1.324) | (0.912) | (1.201) | (1.304) | (2.985) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A6: Regressions Controlling for Counterfactual, 25 percent looser definition of "coder"

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -1.655*** | -1.971*** | -1.574*** | -2.942*** | -3.408*** | -2.536*** | -0.170 | -0.325 | -0.340 |
| | (0.260) | (0.282) | (0.353) | (0.423) | (0.298) | (0.459) | (0.271) | (0.370) | (0.343) |
| Post-GPT*Trend | -1.240* | -0.924 | 1.670** | 0.381 | 0.846 | -0.605 | -2.714* | -2.559* | 3.395*** |
| | (0.652) | (0.757) | (0.646) | (0.773) | (0.705) | (1.438) | (1.374) | (1.534) | (0.687) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A7: Regressions Controlling for Counterfactual, 50 percent stricter definition of "coder"

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -2.162*** | -2.288*** | -2.040*** | -2.976*** | -3.088*** | -2.936*** | 0.0797 | -0.0737 | 0.392 |
| | (0.125) | (0.156) | (0.178) | (0.182) | (0.210) | (0.246) | (0.277) | (0.329) | (0.316) |
| Post-GPT*Trend | -1.795* | -1.669 | -5.116*** | -3.154*** | -3.041*** | -6.996*** | 1.318 | 1.471 | -0.738 |
| | (1.041) | (1.085) | (1.372) | (1.110) | (1.143) | (0.961) | (1.137) | (1.227) | (2.821) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A8: Regressions Controlling for Counterfactual, 50 percent looser definition of "coder"

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -1.360*** | -1.480*** | -1.296** | -3.027*** | -3.368*** | -2.805*** | 0.00567 | 0.00994 | 0.0280 |
| | (0.392) | (0.414) | (0.525) | (0.756) | (0.779) | (0.943) | (0.233) | (0.256) | (0.349) |
| Post-GPT*Trend | -4.231*** | -4.111*** | -6.952*** | -4.622*** | -4.281*** | -7.750*** | -4.033*** | -4.037*** | -6.644*** |
| | (0.866) | (0.896) | (1.003) | (1.148) | (1.135) | (1.071) | (0.867) | (0.930) | (1.353) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A9: Regressions Controlling for Counterfactual, excluding software developers

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | 1.792*** | 1.666*** | 1.794*** | 0.128 | -0.107 | 0.316 | 3.130*** | 3.106*** | 2.987*** |
| | (0.202) | (0.200) | (0.280) | (0.375) | (0.347) | (0.446) | (0.217) | (0.237) | (0.267) |
| Post-GPT*Trend | -2.246*** | -2.120*** | -3.014*** | -2.185*** | -1.951*** | -3.646*** | -2.339*** | -2.315*** | -2.665*** |
| | (0.437) | (0.421) | (0.683) | (0.650) | (0.615) | (0.688) | (0.398) | (0.410) | (0.826) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A10: Regressions Controlling for Counterfactual, using all workers to compute counterfactual industry growth

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -0.962*** | -1.184*** | -0.849* | -2.359*** | -2.797*** | -1.967*** | 0.616*** | 0.579*** | 0.533* |
| | (0.331) | (0.300) | (0.437) | (0.642) | (0.568) | (0.743) | (0.206) | (0.211) | (0.287) |
| Post-GPT*Trend | -3.211*** | -2.989*** | -4.405*** | -3.590*** | -3.152*** | -6.000*** | -3.002*** | -2.965*** | -3.382*** |
| | (0.632) | (0.594) | (0.845) | (1.080) | (1.006) | (1.172) | (0.410) | (0.416) | (0.811) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A11: Regressions Controlling for Counterfactual, using employment shares instead of levels

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -1.325*** | -1.542*** | -1.202*** | -2.395*** | -2.815*** | -2.016*** | -0.0149 | -0.0569 | -0.0667 |
| | (0.344) | (0.315) | (0.456) | (0.635) | (0.573) | (0.738) | (0.208) | (0.202) | (0.312) |
| Post-GPT*Trend | -3.683*** | -3.467*** | -4.673*** | -3.696*** | -3.276*** | -5.848*** | -3.723*** | -3.680*** | -3.925*** |
| | (0.605) | (0.568) | (0.794) | (1.039) | (0.974) | (1.183) | (0.374) | (0.373) | (0.686) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A12: Regressions Controlling for Counterfactual, monthly industry respondent threshold of zero

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | 0.0200 | -0.181 | 0.0181 | -1.429** | -1.817*** | -1.226* | 1.665*** | 1.627*** | 1.522*** |
| | (0.296) | (0.301) | (0.397) | (0.547) | (0.515) | (0.665) | (0.234) | (0.255) | (0.291) |
| Post-GPT*Trend | -3.130*** | -2.929*** | -4.236*** | -3.465*** | -3.077*** | -5.352*** | -2.952*** | -2.915*** | -3.540*** |
| | (0.586) | (0.554) | (0.809) | (0.959) | (0.888) | (1.055) | (0.429) | (0.440) | (0.827) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A13: Regressions Controlling for Counterfactual, monthly industry respondent threshold of 50

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -1.800*** | -1.999*** | -1.662*** | -3.365*** | -3.778*** | -2.891*** | -0.156 | -0.174 | -0.256 |
| | (0.346) | (0.333) | (0.453) | (0.681) | (0.631) | (0.752) | (0.203) | (0.205) | (0.305) |
| Post-GPT*Trend | -2.306*** | -2.108*** | -4.136*** | -1.831 | -1.418 | -5.475*** | -2.564*** | -2.546*** | -2.644*** |
| | (0.762) | (0.752) | (0.753) | (1.237) | (1.201) | (1.169) | (0.427) | (0.439) | (0.796) |
| Observations | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A14: Regressions Controlling for Counterfactual, sample ending in 2024 with pre-NAICS 2022 codes

| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | 0.425* | 0.341* | 0.491 | -0.644 | -0.949 | 0.120 | 0.616*** | 0.579*** | 0.533* |
| | (0.216) | (0.185) | (0.309) | (0.805) | (0.700) | (0.802) | (0.206) | (0.211) | (0.287) |
| Post-GPT*Trend | -4.301*** | -4.216*** | -5.299*** | -9.133*** | -8.828*** | -13.09*** | -3.016*** | -2.978*** | -3.410*** |
| | (0.583) | (0.580) | (0.805) | (1.773) | (1.783) | (1.430) | (0.414) | (0.420) | (0.818) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A15: Regressions Controlling for Counterfactual, dropping NAICS 5415 (computer systems design and related services)

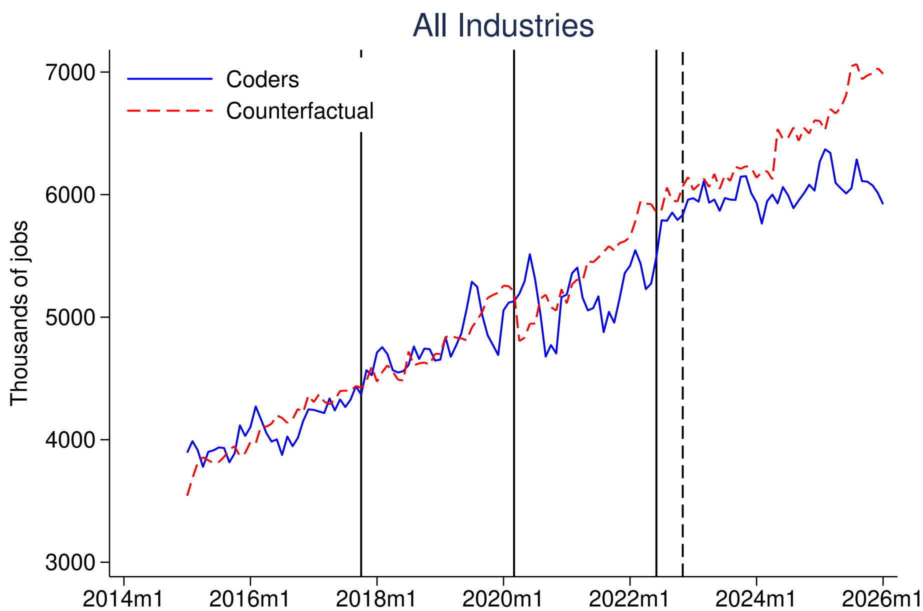
| | All industries | | | Coding intensive industries | | | Non-coding intensive industries | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Trend | -0.922*** | -1.139*** | -0.833* | -2.310*** | -2.743*** | -1.957*** | 0.645*** | 0.613*** | 0.548* |
| | (0.318) | (0.293) | (0.423) | (0.619) | (0.550) | (0.726) | (0.209) | (0.217) | (0.283) |
| Post-GPT*Trend | -3.415*** | -3.197*** | -4.894*** | -3.769*** | -3.336*** | -6.504*** | -3.186*** | -3.153*** | -3.797*** |
| | (0.657) | (0.623) | (0.831) | (1.101) | (1.030) | (1.145) | (0.434) | (0.442) | (0.809) |
| Observations | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 | 133 |
| Sample | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 | All | No Covid | No 2022/3 |

Standard errors in parentheses

* p<.10, ** p<.05, *** p<.01

*Note:* Coefficients are annualized log points, standard errors are Newey-West. Dependent variable is difference between monthly (log) coder employment and the counterfactual (log) employment series. Post-GPT period starts in November 2022.

Table A16: Regressions Controlling for Counterfactual, using smoothed Coglianese et al. (2025) weights




*Note:* Employment level of coder occupations based on O*NET programming skill. Vertical lines show structural break dates estimated via Bai-Perron tests. The dashed line shows ChatGPT release date (November 2022).
*Source:* O*NET, CPS, authors' calculations

Figure A1: Bai-Perron Breakpoints