# Examining the Economy on an Occupational Level: A Baseline-Sensitivity Warning for AI-Employment Research

Caleb B. Eng

July 19, 2026

**Abstract**

Using BLS Occupational Employment and Wage Statistics (OEWS) data on over 860 occupations, I document a measurement problem that affects AI-employment research: any study that measures recent growth against a pre-period baseline spanning 2019-22 is vulnerable to a distortion caused by the COVID-19 shock and recovery rather than by whatever the researcher is actually trying to capture. I use the ongoing debate over AI's employment effects as a concrete test case, but the underlying issue belongs to anyone comparing post-2022 growth to a pandemic-era baseline. I show the problem with three independent pieces of evidence: (1) an own-trend baseline and an independent external-forecast baseline, applied to the same occupations, agree only weakly and can produce opposite conclusions about identical cases; (2) decomposing the resulting correlation algebraically shows it is dominated by how the measure is constructed, not by genuine mean reversion between periods, meaning the pattern would appear even if the underlying growth rates were unrelated; and (3) splitting employment growth into sub-periods shows AI exposure correlates with growth specifically during the COVID shock and recovery windows, but has no detectable relationship with growth either before COVID or in the period the research question is actually about. Separately, I show AI exposure is robustly, positively associated with pre-ChatGPT wage and that this same wage variable independently predicts how exposed an occupation was to the pandemic shock, tying the two results together: a pattern that looks AI-specific is at least partly a wage and occupation-type divide that predates AI entirely.

# 1 Introduction

Since the release of ChatGPT in November 2022, AI has received extensive media coverage leading economists to wonder what effects it will have on the economy and which occupations will be most affected. In this paper, I use BLS data from the Occupational Employment and Wage Statistics (OEWS) tables to examine over 860 detailed occupations. I make a methodological research note explaining how a common way of measuring the effect of AI, namely comparing post-2022 growth to a pre-ChatGPT baseline, is unreliable in ways easy to miss. I do not make a claim about AI's employment effects. Because OEWS captures annual employment *headcounts* rather than vacancies, hours, or task content, it is well suited to the measurement question this paper asks but not necessarily to detecting the earliest labor-market effects of generative AI, which more plausibly show up first in postings or task composition before they show up in annual headcounts.

Because 2019-22 growth in AI-exposed occupations was elevated *relative to* less-exposed occupations for a reason that has nothing to do with AI, mistakes arise naturally. The mechanism, developed in full in Section 4.3 and Section 5, is differential COVID-19 insulation rather than a boom that later corrected, which may still have happened. Office- and knowledge-work occupations, which, as I will show, are also the higher-wage, higher-AI-exposure occupations, took a comparatively small hit in the initial 2020 shock, while in-person service and physical-labor occupations crashed harder and were still working their way back toward a pre-pandemic trend well into 2022. Over the full 2019-22 window, AI-exposed occupations therefore show elevated growth relative to less-exposed ones at least partially because they largely avoided the pandemic-era dip that everyone else had to recover from. Independently, the broader 2021-22 labor market was also historically tight[1] which is a separate reason any measure anchored near 2022, including a fixed post-ChatGPT baseline such as Brynjolfsson et al. (2025)'s, should be treated cautiously; I return to this point in Section 2.

Whatever the exact mix of mechanisms behind an elevated pre-period baseline, subtracting an inflated pre-period growth rate mechanically produces a negative-looking Bend measure even absent any true reversal (Section 4), and a period-by-period decomposition (Figure 3) shows the exposure-growth correlation is concentrated exactly in the COVID shock and recovery windows and vanishes in the post-ChatGPT window itself.

I go further and show that wage is positively correlated with growth not only during the COVID era but in the 2022-25 period as well, and although AI exposure is itself correlated with wage, its raw correlation with 2022-25 growth is approximately zero. Adjusting for wage, a simple regression suggests exposure may reduce employment growth, though not at a conventionally significant level; the result does not survive the addition of occupational-category (2-digit SOC major-group) fixed effects, the same control that already weakens this paper's other exposure-growth findings.

Brynjolfsson et al. (2025), currently the most prominent large-scale empirical study of AI's employment effects, independently encounters a version of the same problem using the identical exposure measure, and flags it in a sentence, addressing it only indirectly via robustness checks that never quantify the pre-trend. I want to be precise about what this paper claims in relation to that work. Their main design differs from mine: they use high-frequency worker-level ADP payroll data, compare age groups within the same firm, and normalize employment to a single post-ChatGPT date rather than differencing two growth rates, avoiding some of what I document below. I am not claiming to have shown that their finding of declining employment

[1]The JOLTS vacancy-to-unemployment ratio peaked at roughly 2.0 job openings per unemployed worker in March 2022, about 65% above its already-tight 2019 average of roughly 1.2 (FRED).

for young workers in AI-exposed occupations is a mean-reversion artifact. Doing so would require showing something specific about age-differentiated hiring within firms before 2022, which is outside what an occupation-level OEWS panel can speak to. Rather, I document, in an independent dataset, precisely the kind of pandemic-baseline contamination that they flag as an open concern for this exact exposure measure, and I show its magnitude and mechanism in a way their robustness checks do not.

# 2 Related Literature

AI exposure scores in this paper come from Eloundou et al. (2024), who use GPT-4 (validated against human labeling) to rate how capable AI is of completing the tasks that make up each O*NET-defined occupation. For reasons discussed later, I primarily use their model-rated $\beta$ measure, which captures exposure once AI is combined with additional tooling (e.g., image generation, code execution). I also refer to their model-rated $\alpha$, which captures exposure over a bare chat interface alone.

Brynjolfsson et al. (2025) use high-frequency ADP payroll data to show that early-career workers (ages 22-25) in AI-exposed occupations have experienced substantial relative employment declines since late 2022. Their main specification is an event study in employment levels normalized to October 2022, with firm-time fixed effects. Even so, they explicitly flag the concern: *"One alternative confounder that would not be controlled for with firm-time effects is that... workers with high AI exposure were excessively hired after the COVID-19 pandemic, leading to a subsequent contraction in their hiring"* (p. 12). When they extend their sample back to 2018 as a placebo check, they find: *"for the Eloundou et al. (2024) measures the most exposed quintile had slower employment growth starting around 2020... This is not the case for the Anthropic exposure measures"* (p. 14). They do not quantify this pre-trend or explain its mechanism; they simply note that it does not appear when using the Anthropic exposure measure instead.

Their own introductory summary of findings states that *"the AI exposure taxonomy did not meaningfully predict employment outcomes for young workers further back in time, before the widespread use of LLMs, including during the unemployment spike driven by the COVID-19 pandemic"* (p. 4). This is in tension with their own concession, discussed above, that for the Eloundou et al. (2024) measure specifically, the most exposed quintile had slower employment growth starting around 2020 (p. 14). Their main balanced panel, which requires continuous firm records from January 2021 onward, cannot speak to 2019-2020 directly; the only sample that can is the smaller, noisier 2018-extension where this contradiction appears.

This concern should not be waved off as easily as their brief flag suggests. Normalizing employment to a single October 2022 baseline, as their main specification does, is not obviously safer than the Bend19-22 measure I critique below: both treat 2022 employment levels as a "normal" reference point, and both are therefore vulnerable to whatever made 2022 an unusual year in exactly the occupations under study.

This raises an obvious question: why would AI-exposed occupations specifically have experienced unusual 2022 overhiring, rather than the economy generally? There are two plausible, AI-unrelated answers. First, as the Introduction notes, the 2021-22 labor market was tight economy-wide (JOLTS vacancy-to-unemployment ratio near 2.0), and that tightness was not evenly distributed. High-wage, office-based occupations, which are also disproportionately the high-AI-exposure occupations in this dataset (Section 5), compete in a labor market where hoarding scarce skilled workers is comparatively easy to justify. Second, and independently, 2021-22 was the tail end of a genuine pre-ChatGPT technology and digital-economy hiring boom. Remote work, e-commerce, and software investment expanded sharply in exactly the tech-adjacent, white-collar occupations that later score high on AI exposure, for reasons that predate ChatGPT's release by years and have nothing to do with AI.

I want to be clear that this overhiring hypothesis is offered as a reason their own fixed-2022 baseline deserves scrutiny, not as a claim about the mechanism documented in this paper: Section 4.3's period decomposition points to differential COVID-19 insulation (exposed occupations barely dipping, rather than dipping and then surging) as this paper's own evidence, not a boom-and-correction story of the kind their footnote describes. In short, the two papers may both be picking up 2019-22 baseline contamination without sharing the same underlying mechanism. Section 4 of this paper is best read as an independent stress test of the same underlying pandemic-baseline problem, not a rebuttal of their specific result. Their design differs enough (individual-level, age-differentiated, within-firm) that an occupation-level aggregate finding cannot mechanically overturn it. What this paper adds is a demonstration, with a mechanism, of exactly how badly this exposure measure's relationship with pre-2022 growth can distort a naive acceleration measure, which is directly relevant to interpreting their footnote even if it does not settle their main result.

The wage-exposure relationship I document in Section 5 has a direct historical predecessor. Frey and Osborne (2017) found that computerization risk during the digital-computerization era was negatively correlated with wage and education. I want to flag a construct difference before treating this as a clean reversal: Frey and Osborne estimate the probability that an occupation's tasks could eventually be fully automated, while Eloundou et al.'s $\beta$ measures the share of tasks for which current LLMs (with tooling) could plausibly reduce completion time. A high-$\beta$ occupation is not necessarily at high risk of being automated away; it may simply be one where AI tools currently help. With that caveat, the empirical fact stands on its own: the occupational incidence of Eloundou-style exposure is the opposite of the occupational incidence of Frey-Osborne-style automation risk. This is a real and useful descriptive fact, not proof that AI will hit the opposite end of the wage distribution from digital-era automation.

Beyond the academic literature, the measurement difficulty this paper documents is also live in financial and popular press coverage. The Economist's "The tech jobs bust is real. Don't blame AI (yet)" (April 2026) makes a parallel argument in a different register, and adds a related wrinkle: even November 2022 may be the wrong reference point for the AI era, since genuinely capable coding tools did not arrive until roughly two years later. Greg Ip's "AI Is Distorting Practically Everything About the Economy" (WSJ, May 2026) makes the broader case that AI investment is currently distorting GDP and productivity measurement generally. This review is not exhaustive: a fuller academic literature on AI and job loss almost certainly exists beyond Brynjolfsson et al. (2025) and the digital-era comparison in Frey and Osborne (2017), and surveying it properly is a task for a future version of this paper rather than something attempted here.

# 3 Data and Measures

The core panel covers 861 detailed BLS occupations (2018 SOC revision) with OEWS employment counts from 1997-2025, matched to Eloundou et al. (2024) exposure scores and to BLS's own 2021-31 Employment Projections. OEWS is an annual establishment survey of wage-and-salary employment stocks, not vacancies, hours, hiring rates, or task content (see the Introduction for the resulting data-coverage caveat). I define two measures of post-2022 acceleration:

- **Bend19-22** $=\text{Growth}_{22\text{-}25}-\text{Growth}_{19\text{-}22}$: the occupation's 2022-25 annualized log-point growth rate minus its own 2019-22 growth rate.
- **BendPretrend** $=\text{Growth}_{22\text{-}25}-\text{BLSForecast}_{21\text{-}31}$: the same numerator, benchmarked instead against BLS's own annualized growth expectation from the 2021-31 Employment Projections, published in 2022 before ChatGPT's release.

I want to acknowledge upfront that BendPretrend is not a clean counterfactual either. It has a 2021 base year that is itself pandemic-affected; it is a ten-year structural projection, not a short-run forecast, so BLS was not attempting to predict cyclical movement between 2022 and 2025; and its errors may embed BLS's own assumptions about technology adoption, industry reallocation, and much else that has nothing to do with AI specifically. The comparison between the two measures in Section 4 is informative because it demonstrates sensitivity to baseline choice, not because BendPretrend is closer to an identified AI effect.

Across the 861-occupation backbone, 33.0% of the labor force was in a decelerating occupation on the Bend19-22 measure as of 2025, with 65% of occupations posting accelerations within $\pm 5\%$ and 6.2% of employment share missing data due to classification changes (Figure 1). Similarly, for BendPretrend, 31.7% of the labor force was in a negative bend occupation and 87% of occupations were within $\pm 5\%$. AI exposure scores are available for 770 of the 861 occupations; Bend19-22 for 757; BendPretrend for 829.

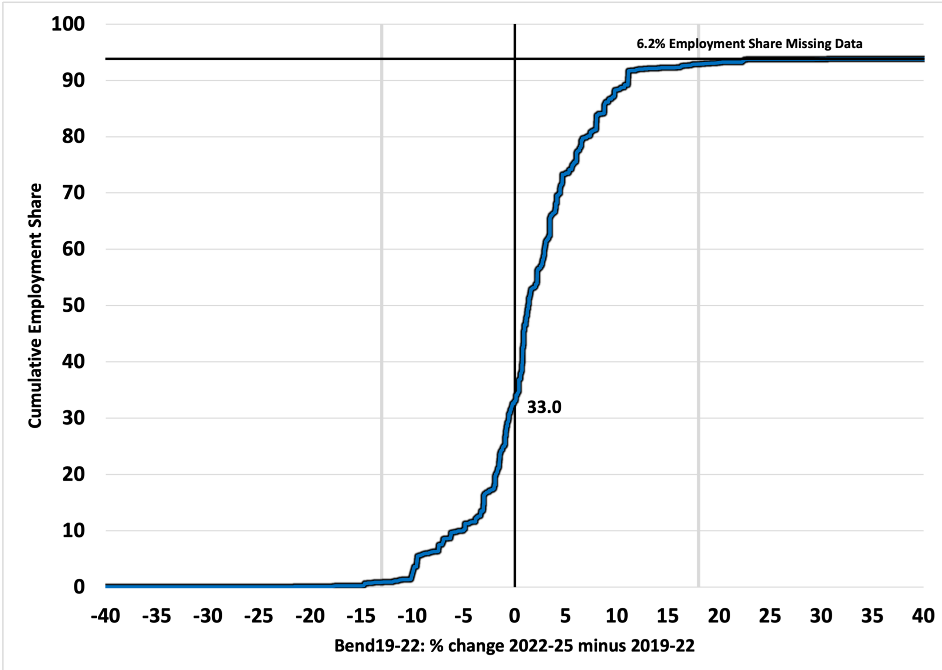


Figure 1: Cumulative distribution of occupations by Bend19-22 acceleration, weighted by 2022 employment share. Most sizable occupations are between -13 and 18 percentage points.

# 4 Result 1: Acceleration Measures Are Highly Sensitive to Baseline Choice

## 4.1 The two measures disagree

Bend19-22 and BendPretrend correlate only moderately with each other.[2] Figure 2 is perhaps a more intuitive way to see the same disagreement than the bare correlation coefficient, but to preserve visual harmony I omit observations with less than 5,000 employees. Using the horizontal and vertical zero-lines to split occupations into quadrants, 32.3% of 2022 national employment sits in the two quadrants where the measures give *opposite-signed* readings: 16.8% in occupations Bend19-22 calls decelerating but BendPretrend calls accelerating, and 15.4% the reverse. Computer Programmers is the clearest single-occupation illustration: Bend19-22 $= +1.45$ (reads as mild acceleration) while BendPretrend $= -11.06$ (one of the largest misses versus BLS's own forecast in the entire dataset). Telemarketers ($-5.6$ / $-14.7$) and Statisticians ($+6.0$ / $-4.8$) are red-dotted alongside it as two further canonically AI-exposed occupations with large discrepancies.[3] The same occupation, in the same dataset, can produce opposite headline readings depending on the chosen counterfactual.

[2]Pearson $r = 0.64$, Spearman $\rho = 0.59$ ($n = 691$, restricted to occupations with at least 5,000 workers in 2025)

[3]Note that three other canonically AI-exposed occupations — Data Scientists, Web Developers, and Software Developers — are excluded from this comparison because Bend19-22 is not computable for them in this data.

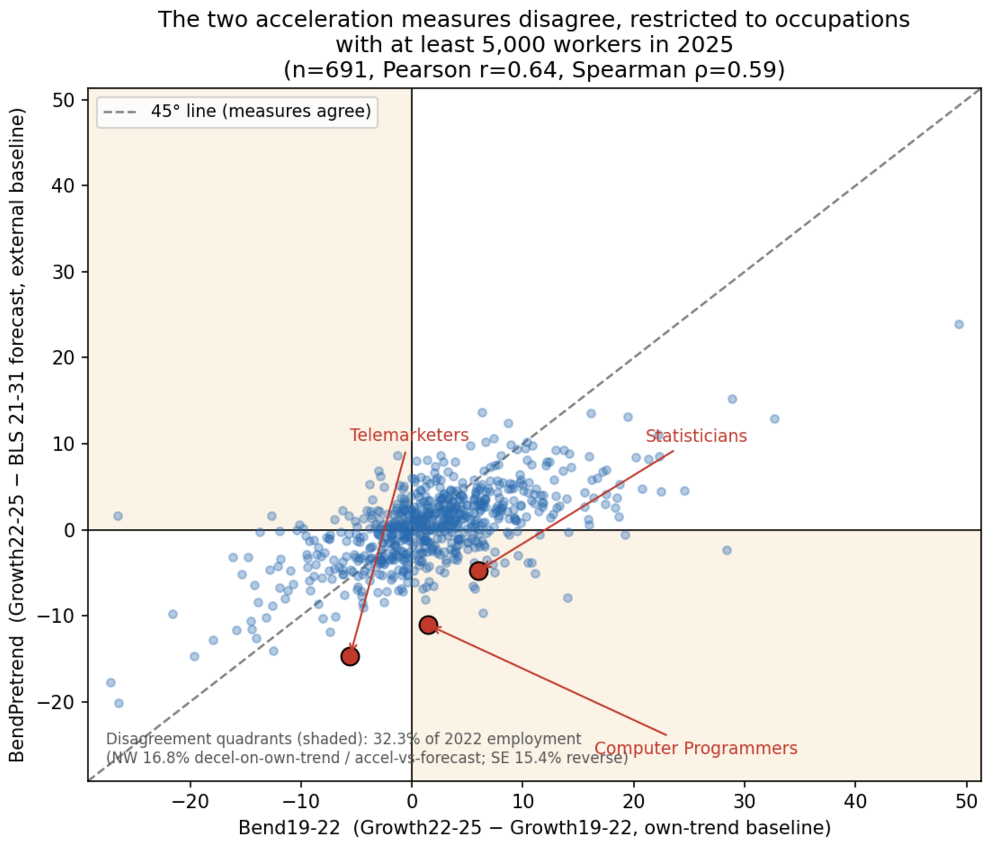


Figure 2: Bend19-22 vs. BendPretrend, restricted to occupations with at least 5,000 workers in 2025. The 45° line marks where the two measures would agree; shaded regions mark the two disagreement quadrants.

## 4.2 Mechanism, part 1: how much of this is pure algebra?

Bend19-22 correlates $-0.746$ with the very growth rate (Growth19-22) it subtracts. Before treating that as evidence of a real economic relationship, it is worth asking how much of it is guaranteed by construction. For any measure Bend $= g_{\text{post}} - g_{\text{pre}}$,

$$\text{Cov}(\text{Bend}, g_{\text{pre}}) = \text{Cov}(g_{\text{post}}, g_{\text{pre}}) - \text{Var}(g_{\text{pre}}),$$

so a negative correlation between Bend and $g_{\text{pre}}$ is partly guaranteed regardless of whether the two growth rates are related at all, because $\text{Var}(g_{\text{pre}})$ enters with a negative sign automatically. I compute both terms directly. If $\text{Cov}(g_{\text{post}}, g_{\text{pre}})$ were exactly zero — i.e., pre- and post-period growth were completely unrelated, algebra alone would already produce $\text{corr}(\text{Bend19-22}, g_{\text{pre}}) = -0.813$. The actual observed correlation is $-0.746$, which is *less* negative than the pure-algebra floor, because $\text{corr}(g_{\text{post}}, g_{\text{pre}})$ is not negative at all, rather it is $+0.100$. Occupations that grew faster 2019-22 grew mildly *faster*, not slower, 2022-25. There is no genuine mean reversion in the classical sense here; what looks like a strong mean-reversion signal in Bend19-22 is overwhelmingly a mechanical consequence of subtracting a noisy pre-period term from itself, only partially offset by mild positive persistence pulling the correlation back toward zero. Thus, *any* variable that happens to correlate with Growth19-22, for COVID-related reasons or any other reason, will show up with an amplified, sign-flipped relationship to Bend19-22 by construction, independent of whether that variable has any true relationship with 2022-25 growth. Section 4.3 shows this is close to exactly what happens with $\beta$.

## 4.3 Mechanism, part 2: a period decomposition

If $\beta$'s relationship with Growth19-22 reflects a genuine, stable feature of AI-exposed occupations rather than a pandemic-specific artifact, it should also show up before COVID and persist into the post-ChatGPT period. I split employment growth into four sub-periods — 2015-19 (ordinary pretrend), 2019-20 (the initial COVID shock, one year), 2020-22 (recovery), and 2022-25 (post-ChatGPT) — and test whether $\beta$ predicts each one separately, on a fixed sample of 656 occupations with complete data in every sub-period (Figure 3).

The pattern is stark. $\beta$ has no detectable relationship with ordinary pre-pandemic growth: exposed and non-exposed occupations were on statistically indistinguishable trends before 2020. It then correlates significantly with growth during the COVID shock itself and, more weakly, during the 2020-22 recovery. It has no detectable relationship with growth in the post-ChatGPT period, 2022-25. Put plainly, the two mechanism results fit together like this: the algebra above shows that *any* variable correlated with 2019-22 growth, for any reason, gets picked up and sign-flipped by the Bend19-22 subtraction, whether or not it has anything to do with 2022-25 growth. Figure 3 shows $\beta$ does exactly that and is correlated with growth only in the window the subtraction inherits from (2019-22, especially the COVID shock itself) and uncorrelated with growth in the window the measure is nominally trying to isolate (2022-25). Together, the two results explain why Bend19-22 looks like an AI deceleration effect

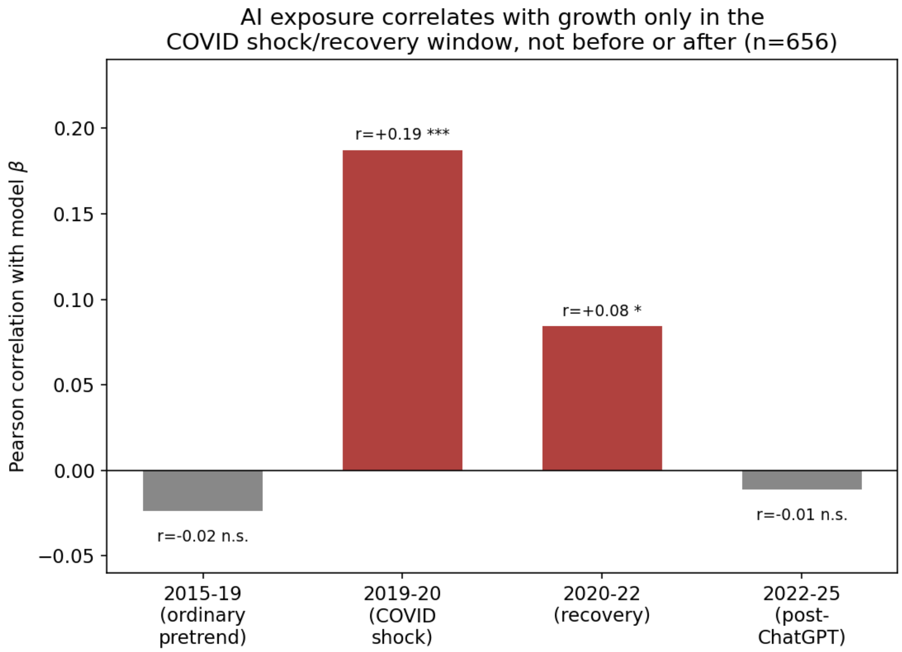


Figure 3: Pearson correlation between model $\beta$ and employment growth in each sub-period, fixed sample ($n = 656$).

without being one. Section 5 takes up directly whether this COVID-window correlation is itself explained by wage.

## 4.4 Does AI exposure survive regression controls?

Sections 4.2-.3 already showed that a naive pre/post-2022 comparison is unreliable by construction and that $\beta$'s apparent relationship with it is a pandemic-window artifact. It is still worth asking whether $\beta$ survives as a predictor once the outcome is regressed on directly rather than just decomposed by period, because a reader could reasonably ask whether the decomposition evidence and a standard regression tell the same story, which they do. I focus on $\beta$ alone and drop $\alpha$ from this table entirely. As Section 5 shows, $\alpha$'s own association is not stable enough on its own to interpret cleanly once $\beta$ is accounted for, so I defer that discussion here and keep this table focused on $\beta$. Note that other specifications using $\alpha$ do not change the conclusion below, only its presentation.

Table 1 reports $\beta$'s coefficient predicting each Bend measure, pooled and with major occupational-group (2-digit SOC) category fixed effects added. The two Bend measures disagree even in sign: $\beta$ is significantly negative for Bend19-22 but statistically indistinguishable from zero for BendPretrend. Neither survives the addition of category controls. Bend19-22's coefficient loses significance once occupational category is held fixed, and BendPretrend's is never significant at the 5% level, though the category-FE specification comes close ($p = .093$). This is exactly what Sections 4.2-.3 would predict: if the relationship is a pandemic-window construction artifact rather than a stable feature of the data, it should be fragile to both baseline choice and to controlling for the broad occupational categories that the COVID shock hit unevenly. As this simple regression shows, it is.

Table 1: $\beta$-alone coefficient predicting each Bend measure, pooled vs. category fixed effects (HC1 robust SEs).

| | Pooled | +Category FE |
|---|---|---|
| Bend19-22 ($n = 707$) | $-5.10^{***}$ ($< .001$) | $-5.04$ (.122) |
| BendPretrend ($n = 770$) | $-0.19$ (.832) | $-2.85$ (.093) |

## 4.5 Robustness check: a multi-point baseline-distance test

Table 2 extends the test to five baseline start years, holding the post-period fixed at 2022-25 and the occupation sample fixed ($n = 657$[4]; plotted in Figure 4). The result is not a smooth monotonic decay: the correlation is roughly flat across the three pure-pretrend starting points, then rises once the window begins to overlap the COVID/recovery period. This matches the period-decomposition finding above. Contamination is specific to windows touching 2020-22, not a generic function of how far back the baseline reaches.

Table 2: Correlation between $\beta$ and raw pre-period growth, by pre-period start year ($n = 657$ throughout).

| Pre-period start year | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|
| Pearson $r$ | 0.080 | 0.070 | 0.068 | 0.097 | 0.154 |

## 4.6 Which occupations deviate most from the pretrend baseline

Table 3 lists the occupations with the largest negative BendPretrend values (restricted to 2025 employment $\geq 7{,}500$). Given the caveats about BendPretrend's own limitations (Section 3), I present this as a list of occupations whose 2022-25 growth fell furthest short of BLS's 2021-vintage forecast, not as a validated ranking of AI impact. The deviations could reflect forecast error, industry reallocation, interest-rate effects, or genuine AI exposure, and this dataset cannot cleanly separate those effects.

Table 4 buckets the same 25 occupations into groups of five and shows how much of the economy this list actually represents. It is small. All 25 occupations together account for 1,331,710 jobs in 2025 — 0.86% of the 155.5 million total employment — against an emp-weighted average wage ($75,618) not far above the rest of the economy's ($69,628). Whatever these occupations' 2022-25 shortfalls reflect — genuine AI exposure, industry-specific factors,

[4]One more than the $n = 656$ sample in Figure 3: this test does not require a standalone 2020 data point, unlike the period decomposition, which needs 2020 to separate the 2019-20 and 2020-22 sub-periods.

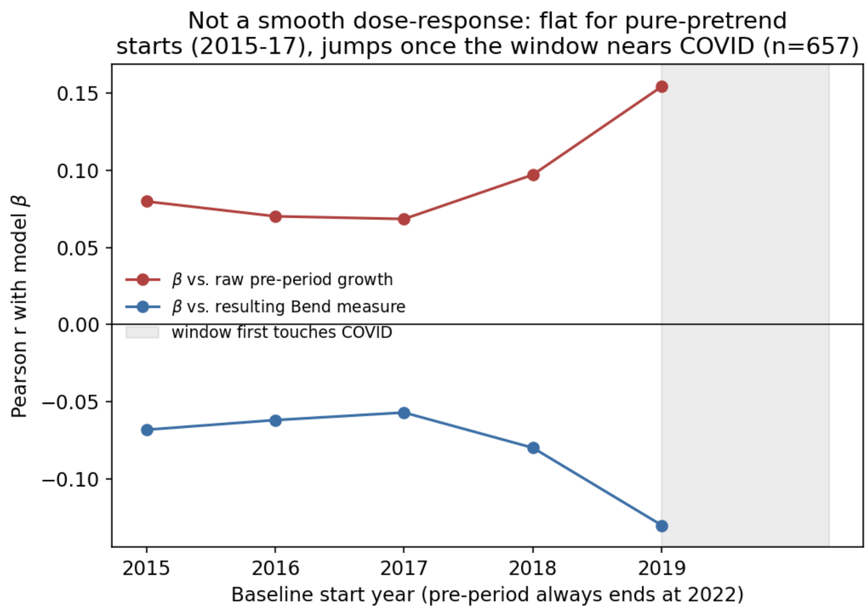


Figure 4: Correlation between $\beta$ and pre-period growth (red) and the resulting Bend measure (blue), for five different pre-period start years, same 657-occupation sample throughout.

Table 3: Occupations with the largest negative pretrend deviations (2025 employment $\geq$ 7,500).

| Occupation | Pretrend Deviation | 19-22 Acceleration (Bend19-22) | Exposure Score ($\beta$) |
|---|---|---|---|
| Special Effects Artists and Animators | -20.2% | -26.4% | 0.525 |
| Telemarketers | -14.7% | -5.6% | 0.806 |
| Broadcast Technicians | -14.7% | -19.6% | 0.524 |
| Loan Interviewers and Clerks | -12.8% | -17.9% | 0.594 |
| Postsecondary Teachers, All Other | -12.6% | -14.1% | n/a |
| Order Clerks | -11.9% | -7.4% | 0.714 |
| Engine and Other Machine Assemblers | -11.7% | -15.8% | 0.095 |
| Medical Appliance Technicians | -11.5% | -14.4% | 0.138 |
| Computer Programmers | -11.1% | 1.5% | 0.950 |
| Coin, Vending, and Amusement Machine Servicers and Repairers | -10.6% | -14.5% | 0.321 |
| Web Developers | -10.4% | n/a | 0.933 |
| Orthopedic Surgeons, Except Pediatric | -10.3% | n/a | 0.333 |
| Plasterers and Stucco Masons | -10.3% | -8.1% | 0.077 |
| Fiberglass Laminators and Fabricators | -10.2% | -13.1% | 0.021 |
| Upholsterers | -10.1% | -6.9% | 0.191 |
| Counselors, All Other | -9.8% | -21.6% | n/a |
| Credit Authorizers, Checkers, and Clerks | -9.7% | 6.4% | 0.650 |
| Boilermakers | -9.0% | -4.4% | 0.030 |
| Film and Video Editors | -8.8% | -12.6% | 0.528 |
| Sales and Related Workers, All Other | -8.6% | -8.5% | n/a |
| Metal Workers and Plastic Workers, All Other | -8.6% | -4.4% | n/a |
| Special Education Teachers, All Other | -8.4% | -13.8% | 0.475 |
| Switchboard Operators, Including Answering Service | -8.1% | 1.2% | 0.629 |
| Loan Officers | -8.0% | -11.5% | 0.591 |
| Grinding and Polishing Workers, Hand | -7.9% | 14.0% | 0.036 |

or forecast error — the "AI threat" visible in this particular ranking, so far, is quantitatively tiny relative to the labor force as a whole.

Table 4: Pretrend-deviation table (Table 3), bucketed into groups of 5. Avg. Wage is 2025 mean wage, employment-weighted within each group.

| Rows | Avg. Pretrend Dev. | Avg. Bend19-22 | Avg. $\beta$ | Total 2025 Emp. | Emp. Share | Avg. Wage |
|---|---|---|---|---|---|---|
| 1-5 | -15.0% | -16.7% | 0.612 | 414,140 | 0.27% | $67,868 |
| 6-10 | -11.4% | -10.1% | 0.444 | 239,120 | 0.15% | $71,931 |
| 11-15 | -10.3% | -9.4% | 0.311 | 139,910 | 0.09% | $108,304 |
| 16-20 | -9.2% | -8.1% | 0.403 | 169,590 | 0.11% | $63,424 |
| 21-25 | -8.2% | -2.9% | 0.433 | 368,950 | 0.24% | $79,916 |
| All 25 | -10.8% | -9.5% | 0.436 | 1,331,710 | 0.86% | $75,618 |
| All other (836 occ.) | — | — | — | 154,150,130 | 99.14% | $69,628 |

# 5 Result 2: AI Exposure and Wages

Table 5 and Figure 5 show the relationship between model $\beta$ and 2019 mean occupational wage chosen so the outcome cannot already reflect an AI-era response. The raw correlation is positive and relatively strong. In a joint regression of log wage on $\alpha$ and $\beta$ together, $\alpha$'s own association flips negative while $\beta$'s remains strongly positive.

Table 5: Log(2019 wage) regressed on AI exposure (HC1 robust SEs, $n = 704$).

| | Coefficient | $p$-value |
|---|---|---|
| Intercept | 10.587 | $< 0.001$ |
| $\alpha$ (raw substitutability) | $-1.830$ | $< 0.001$ |
| $\beta$ (tooling-dependent exposure) | $+1.841$ | $< 0.001$ |
| $R^2$ | 0.384 | |

Figure 5 is not driven by a handful of extreme points: a dense cluster of dark, overlapping dots sits at $\beta \approx 0.35$-$0.50$ toward the high-wage end of the distribution, which raises the question of whether the full-sample log-linear fit is being pulled down (understated) by that cluster's internal scatter, which it is. Excluding the 148 occupations with $\beta > 0.5$ strengthens both the raw correlation[5] and the joint-regression fit.[6] This high-$\beta$ cluster is not a sparse, low-employment tail that featured a handful of small occupations with noisy wage estimates dragging down the fit. Occupations with $\beta > 0.5$ account for 148 of 704 occupations (21.0%) but a full 25.6% of 2019 national employment, and their median employment size (70,485) is actually *larger* than the full sample's median (45,910). The understatement is real, but it reflects genuine wage heterogeneity within a substantively large slice of the workforce; these occupations should not be considered outliers.

[5] Pearson $r$ rises from 0.37 to 0.56

[6] $R^2$ from 0.38 to 0.41; $\alpha$'s own coefficient becomes more negative, $-1.82 \to -2.14$

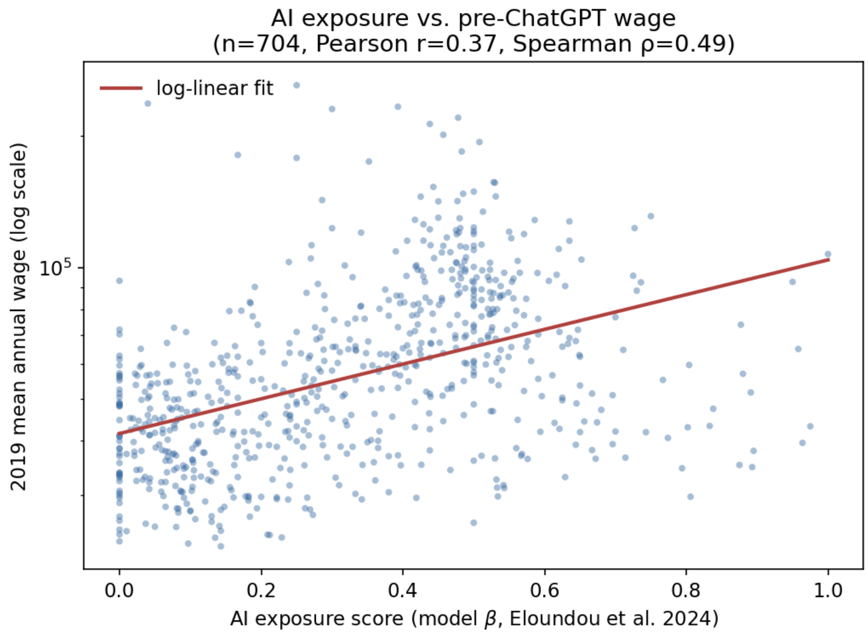


Figure 5: AI exposure (model $\beta$) vs. 2019 mean occupational wage, log scale, with log-linear fit.

The joint regression's sign flip on $\alpha$ also deserves a substantive reading rather than being left as an uninterpreted coefficient. $\alpha$ alone is essentially uncorrelated with 2019 wage[7] meaning raw substitutability by itself is not a proxy for low-wage work, so this is not simply Frey-Osborne-style automation risk reappearing under a new name. What happens once $\beta$ is held fixed is different: the remaining variation in $\alpha$ is disproportionately drawn from occupations where a bare AI chat interface, without additional tooling, can already handle a meaningful share of tasks. High-$\alpha$, comparatively low-$\beta$ occupations in this sample include Bridge and Lock Tenders, Couriers and Messengers, Correctional Officers and Jailers, Bus Drivers, and Occupational Therapy Aides, whose median wage (\$45,830) runs about 15% below the full sample's median (\$53,630). These occupations have a routine, script-following profile, closer to what Frey-Osborne-style automation-risk measures were built to capture than to AI-native knowledge work. Read this way, $\alpha$ net of $\beta$ isolates something closer to the routine/lower-skill sub-component of exposure, while $\beta$ captures the tooling-dependent, higher-skill complement layered on top of it; the joint regression's opposite signs are likely two different slices of the same underlying exposure measure pulling in different wage directions, not a statistical artifact to be explained away.

[7]Pearson $r = 0.01$, $p = 0.72$

## 5.1 Tying it together: does wage explain the COVID-window correlation?

Section 4.3 showed that $\beta$'s relationship with Bend19-22 is concentrated in the COVID shock and recovery windows. I test directly whether wage, which is already strongly related to AI exposure, itself predicts COVID-shock severity, and whether $\beta$'s relationship with the 2019-20 shock survives controlling for wage. Table 6 summarizes the result. Wage predicts the shock more strongly than $\beta$ does on its own, and in a joint regression $\beta$'s own coefficient becomes statistically indistinguishable from zero while wage remains highly significant. This is consistent with $\beta$'s apparent COVID-window correlation being substantially a proxy effect for the same wage/occupation-type divide (office and professional work insulated relative to in-person service and physical work) documented above, rather than a distinct, AI-specific pattern.

Table 6: 2019-20 employment growth on log(2019 wage) and $\beta$ ($n = 704$, HC1 robust SEs).

| | Univariate $r$ (or coef.) | $p$-value |
|---|---|---|
| log(2019 wage), alone | $r = +0.32$ | $< 10^{-5}$ |
| $\beta$, alone | $r = +0.19$ | $< 0.0001$ |
| *Joint regression (2019-20 growth $\sim$ log wage + $\beta$, $R^2 = 0.11$)* | | |
| log(2019 wage) | highly significant | $< 10^{-12}$ |
| $\beta$ | not distinguishable from zero | 0.18 |

Figure 6 extends this comparison to all four sub-periods, mirroring the $\beta$ decomposition in Figure 3.[8] The two do not tell quite the same story, and the difference is informative. Wage correlates with growth in the COVID shock and recovery windows just as $\beta$ does, but unlike $\beta$, whose correlation with growth vanishes in 2022-25[9], wage remains a significant, positive predictor of growth in the post-ChatGPT period too, and even carries a small, marginally significant correlation in the ordinary 2015-19 pretrend that $\beta$ does not. These findings sharpen, rather than undercut, the paper's argument: wage is a general, persistent predictor of occupational growth across every period examined, while $\beta$'s predictive power is specifically confined to the COVID-window years and disappears exactly where a real AI effect would need to show up. The fact that $\beta$ tracks wage during the pandemic but not before or after is more consistent with $\beta$ borrowing wage's COVID-window signal than with wage and $\beta$ jointly reflecting some deeper shared driver of growth in general.

## 5.2 A direct test in the post-ChatGPT window

As noted, $\beta$'s unconditional correlation with 2022-25 growth is essentially zero ($r = -0.01$). But $\beta$ and wage are themselves correlated (Section 5), so an unconditional null does not by

[8] The COVID-shock correlation here is $r = 0.33$ on the $n = 653$ four-period-complete sample, close to but not identical to Table 6's $r = 0.32$ ($n = 704$); the samples differ because the fixed sample requires complete data across all four sub-periods, not just 2019-20.

[9] $r = -0.01$

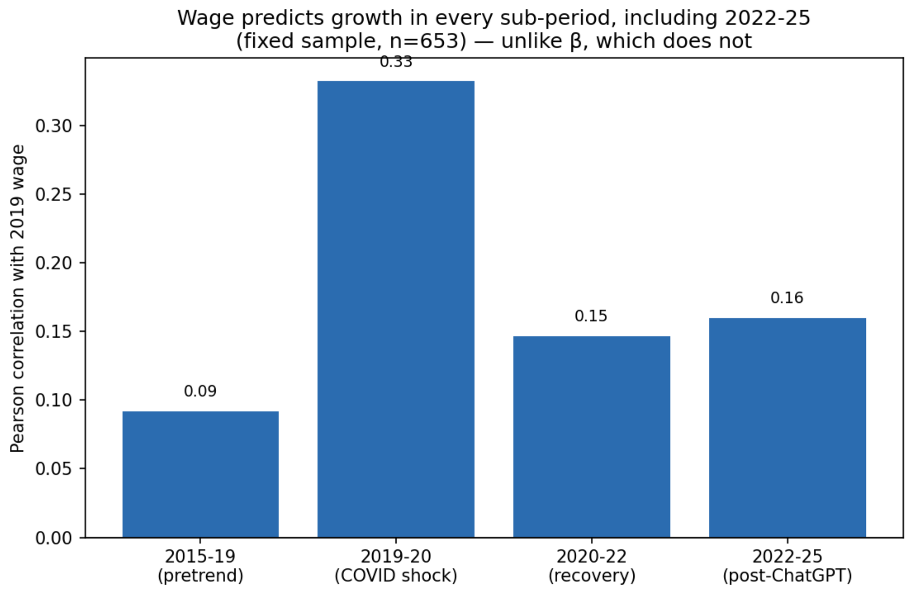


Figure 6: Pearson correlation between 2019 wage and employment growth in each sub-period, same fixed sample as Figure 3, restricted further to occupations with non-missing 2019 wage ($n = 653$).

itself rule out a genuine effect that is masked by wage in the raw correlation. I test this directly by regressing 2022-25 growth on log(2019 wage) and $\beta$ jointly, on the same $n = 704$ sample used throughout this section (Table 7).

Held fixed against wage, $\beta$'s coefficient turns *negative* and is marginally significant ($-2.02$, $p = .071$) — the opposite direction from the near-zero unconditional reading, and consistent with the arithmetic implication that if $\beta$ carried no true post-2022 relationship at all, its raw correlation with growth should still be mildly *positive*, borrowing wage's own positive pull, not flat. Read on its own, this is the closest thing in this paper to a hint of an AI-linked deceleration.

That hint does not survive two natural robustness checks, however. First, adding $\alpha$ to the regression flips $\beta$'s sign back to positive while $\alpha$ itself becomes strongly negative. This is the same $\alpha/\beta$ collinearity documented in the wage regression above (Section 5), where including both components of the same underlying exposure measure can flip either one's sign depending on what the other absorbs. Second, adding occupational-category (2-digit SOC) fixed effects, the same control that already eliminates $\beta$'s relationship with Bend19-22 in Table 1, pushes $\beta$'s coefficient to $p = .110$ in the $\beta$-only specification and renders every coefficient in the table statistically indistinguishable from zero once $\alpha$ is added alongside the fixed effects, while $R^2$ rises from around 0.02-0.04 to roughly 0.18. Category membership, not wage or measured AI exposure, accounts for most of the explainable variation in 2022-25 growth once it is available as a control.

I read this result as consistent with, rather than a challenge to, the paper's central baseline-

Table 7: 2022-25 employment growth regressed on log(2019 wage) and AI exposure, pooled vs. category fixed effects (HC1 robust SEs, $n = 704$).

| | Pooled | +Category FE |
|---|---|---|
| *$\beta$ alone* | | |
| log(2019 wage) | $+2.034^{***}$ | $-0.176$ (.798) |
| $\beta$ | $-2.016$ (.071) | $-3.176$ (.110) |
| $R^2$ | 0.021 | 0.184 |
| *$\alpha$ and $\beta$ jointly* | | |
| log(2019 wage) | $+1.016$ (.067) | $-0.388$ (.576) |
| $\alpha$ | $-7.944^{***}$ | $-2.744$ (.301) |
| $\beta$ | $+2.929$ (.085) | $-1.103$ (.735) |
| $R^2$ | 0.039 | 0.186 |

sensitivity argument. The one specification that hints at an AI-specific effect is also the one specification (pooled, $\beta$ alone) that omits the controls this paper has already shown matter elsewhere. Once either $\alpha$ or category fixed effects are added, the hint disappears. That fragility is itself an instance of the paper's broader point. A result's apparent existence here depends on a specification choice a reasonable researcher could easily have made differently.

# 6 Discussion

The defensible conclusion from Section 4 is narrower than a first pass at this data would suggest. The available occupation-level evidence cannot cleanly distinguish an AI effect from pandemic-era occupational adjustment, and acceleration estimates are highly sensitive to the chosen benchmark, the functional form of the regression, and (per Section 5) a wage/occupation-type composition effect that is at least as plausible an explanation as AI exposure itself. Concretely: imagine three researchers — Amos, Bill, and Cora — each independently handed this same panel and asked whether AI exposure predicts employment deceleration. Amos builds Bend19-22, an own-trend baseline, and confidently reports a significant negative effect: AI-exposed occupations are decelerating ($-5.10$, $p < .001$; Table 1). Bill prefers BendPretrend, an external-forecast baseline, and finds no significant relationship in either direction ($-0.19$, $p = .832$). Cora starts from Amos' own Bend19-22 measure but adds occupational-category fixed effects, a standard robustness check, and watches Amos' finding disappear too ($-5.04$, $p = .122$). All three would defend their own regression as correctly specified, and none of them made an error in their own equation. They would each be reporting a real feature of a fragile measure, with no way to know it from their own regression alone. That fragility, not any single coefficient, is this paper's central finding.

On Brynjolfsson et al. (2025) specifically: this paper does not show their result is mean reversion. What it shows is that the exposure variable underlying their robustness checks has, in an independent occupation-level dataset, a documented and mechanistically explained pandemic-window correlation and so far no post-2022 relationship at all, which contextualizes their brief caveat. It does not demonstrate that their age-differentiated, within-firm result is itself an artifact of the same mechanism. Their design's use of age comparisons within the same firm is a real defense against composition confounds of exactly the kind documented here, but it is not airtight. If young, entry-level workers were disproportionately the flexible margin of any 2022 overhiring, the first hired when a tight labor market pushes firms to expand headcount, and plausibly the first let go when that headcount contracts, then their within-firm, age-differentiated design could still be picking up an age-specific version of the same pandemic-baseline problem, rather than (or in addition to) an AI-specific effect. I want to be precise about what I can and cannot say about this: this paper's data is occupation-level and all-ages, so it cannot see age composition within an occupation or within a firm. This is a real limitation of what an occupation-level OEWS panel can speak to, not a claim that their result is wrong.

# 7 Conclusion and Directions for Future Work

This paper's most defensible finding is a levels-based one, separate from the baseline-sensitivity argument that follows: AI exposure and 2019 wage are robustly, positively related, with $\alpha$ net of $\beta$ reading as closer to the routine/lower-skill sub-component that automation-risk measures were built to capture (Section 5). This comes with the caveat, noted in Section 2, that Eloundou-style exposure and Frey-Osborne-style automation risk are not measuring quite the same construct.

Any measure of 2022-25 employment growth for AI-exposed occupations has to be compared to *something* such as an own pre-period trend, an external forecast, or a fixed reference date, and this paper's central point is that the baseline choice is not a minor implementation detail. Three separate problems compound in the 2019-22 window specifically: differential COVID-19 insulation left less-exposed occupations still recovering well into 2022 (Introduction, Section 4.3); a historically tight 2021-22 labor market makes any baseline anchored near a single 2022 date independently fragile (Section 2); and, for differenced ("Bend") measures specifically, subtracting a pre-period growth rate mechanically amplifies whatever that pre-period happens to correlate with, real or not (Section 4.2). Brynjolfsson et al. (2025)'s fixed-October-2022 baseline is exposed to the first two; this paper's own Bend19-22 measure is exposed to all three.

But the baseline problem is not only a pandemic-era or 2022-specific story, and it is worth closing on the more general version of it. Occupations have always grown at very different secular rates, for reasons that have nothing to do with any particular reference year. Over 2000-2019, Computer Programmers shrank at $-5.15$%/yr while Statisticians grew at $+4.22$%/yr, with Mathematicians ($-0.93$%/yr), Computer and Information Research Scientists ($+0.93$%/yr), and Actuaries ($+2.88$%/yr) spread across the range between them. This is a roughly 9.4-percentage-point spread among occupations that would all plausibly be called AI-adjacent today, driven by nothing but ordinary secular change (outsourcing and abstraction eroding programming as an occupational category, growing statistical/data demand elsewhere).

A fixed single-date baseline like Brynjolfsson et al.'s cannot distinguish an occupation that was already on a structurally different growth path from one newly affected by AI, any more than a differenced measure like Bend19-22 can distinguish a genuine deceleration from an artifact of what its own baseline happened to look like. Both problems point to the conclusion that growth-acceleration claims about AI's employment effect are not safe to report without checking (i) sensitivity to baseline choice, (ii) how much of any observed correlation is mechanical construction rather than signal, (iii) whether the relationship is concentrated specifically in the COVID window rather than present throughout, and (iv) whether it survives controlling for the same wage/occupation-type variables, and the same pre-existing secular growth differences, that predict an occupation's trajectory independent of AI.

The single highest-value extension would be to build pre-determined pandemic-sensitivity controls (2019 wage, education composition, remote-work feasibility where available, industry mix, and each occupation's own pre-2019 secular growth rate) directly into the period-decomposition regressions in Section 4.3 and the wage-COVID-link analysis in Section 5, to test whether $\beta$'s 2019-20 and 2020-22 coefficients are fully explained by these observables or retain an independent residual. It would also be worth attempting a version of the age-differentiated design closer to Brynjolfsson et al.'s own, which this OEWS occupation-level panel cannot support on its own. Finally, narrowing to a small set of canonically AI-exposed occupations — mathematicians, computer programmers, software developers, web developers, data scientists, semiconductor processing technicians, and financial analysts — and tracing each through the full period decomposition individually, in the spirit of the Computer Programmers case study in Section 4.1, would let these general patterns be checked against specific, interpretable cases rather than relying on pooled regressions.

# References

Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen. Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. Working paper, Stanford Digital Economy Lab / Stanford Institute for Economic Policy Research (SIEPR), November 2025.

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308, 2024. doi: 10.1126/science.adj0998.

Carl Benedikt Frey and Michael A. Osborne. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280, 2017.

**AI-Generated Content Disclosure.** This paper was written with assistance from an AI language model (Claude, Anthropic), used as a research and drafting tool throughout data analysis, writing, and revision. The underlying data, empirical results, and analytical judgments are the author's; the AI assisted with code, prose drafting, and editing.