The Labor Hub

Discussion Paper No. 10

LABOR MARKET CONSEQUENCES OF GENERATIVE AI: EARLY EVIDENCE FROM NORWAY

Roberto Iacono and Dennis Facius

Labor Market Consequences of Generative AI: Early Evidence from Norway*

Dennis Facius (EUI)† Roberto Iacono (NTNU) ‡

June 5, 2026

Abstract

Does Generative AI displace early-career workers? We provide population-wide evidence from Norwegian administrative registers, 2015 through March 2025, exploiting the November 2022 release of ChatGPT as an availability shock. Using the within-firm composition difference-in-differences employed in recent work, supplemented with a synthetic difference-in-differences at the occupation level and a firm-level shift-share design, we find no robust evidence of employment displacement among young workers in highly AI-exposed occupations, nor any robust response across other age cohorts or on incumbent labor-market outcomes. While estimated coefficients for young workers are negative, in line with the existing literature, they are small and statistically insignificant. A backdating exercise on the synthetic difference-in-differences yields larger absolute estimates than the actual treatment date across most age bands. This suggests the apparent post-2022 decline reflects, at least in part, pre-existing secular trends rather than a clean AI-period break.

Keywords: *Generative Artificial Intelligence; Large Language Models; Automation; Labor demand.*
JEL Classification: *J23; J24; J31; O33.*

*We are grateful to Francesco Drago, Magnus Lodefalk, Fabrizia Mealli, Tommaso Nannicini, Julian Tiedtke and seminar participants at the EUI Department of Economics (May 2026) for helpful comments. Iacono acknowledges financial support from the Research Council of Norway, project no. 315765 and 358025. The data used in this study relies on individual-level employment records collected through the A-ordningen. The data was extracted and made available for research purposes by Statistics Norway (SSB). The views expressed herein are those of the authors and do not necessarily reflect those of our institutions. All errors are our own.

†Address: European University Institute (EUI), Department of Economics, Via delle Fontanelle 18, 50014 San Domenico di Fiesole (FI), Italy. E-mail: dennis.facius@eui.eu.

‡Address: Norwegian University of Science and Technology (NTNU), Postboks 8900, NO-7491 Trondheim, Norway. E-mail: roberto.iacono@ntnu.no. Iacono is Visiting Senior Fellow at the LSE International Inequalities Institute, CESifo Research Network Affiliate and fellow at the World Inequality Database (WID.World) at PSE.

1 Introduction

Technological change has re-emerged as a central topic in labor economics. While an earlier generation of research focused on industrial robots and routine-biased computerization, the rapid diffusion of Generative Artificial Intelligence (GenAI) in recent years has renewed concerns about how automation reshapes labor markets. Eloundou, Manning, Mishkin, and Rock (2024) estimate that around 19% of the total U.S. workforce could see *at least half* of their tasks affected by large language models (LLMs) alone.

However, the canonical task-based framework (Acemoglu & Restrepo, 2019; Autor, Levy, & Murnane, 2003) predicts that new technologies act on labor markets through partially offsetting channels: *displacement* from the tasks the technology can now perform; *reinstatement* through the creation of new tasks; and a *productivity effect*, whereby falling production costs and rising output raise labor demand for tasks that remain labor-intensive. The net effect on employment and wages therefore remains ambiguous. Applied to the ongoing automation wave, Freund and Mann (2026) argue that workers with moderate overlap with GenAI may benefit from complementarities and productivity gains, while those with high exposure are harmed due to displacement effects. Specifically, returns to social skills are expected to rise relative to analytical skills, and Acemoglu, Autor, and Johnson (2026) identify early-career workers in affected occupations as the most vulnerable group, based on the notion that artificial intelligence cannot yet substitute for experience accumulated by older workers. This paper provides further empirical evidence on the age-gradient hypothesis, based on population-level register data from Norway, with workers aged 22–25 as the focal early-career cohort, in line with existing empirical work.

Early studies on online platforms for freelancers and vacancy postings find substantial negative effects of GenAI on employment and remuneration (Hui, Reshef, & Zhou, 2024; Teeselink, 2025; Teutloff et al., 2025). Using worker-level data from the largest payroll provider in the United States, Brynjolfsson, Chandar, and Chen (2025) document a 13% relative decline in employment of 22–25-year-olds in the most AI-exposed occupations, and Hosseini and Lichtinger (2025) find a corresponding drop in junior hiring at U.S. firms that adopt Generative AI.

More recent research is starting to use linked administrative records, and the resulting picture is notably mixed. Lodefalk, Löthman, Koch, and Engberg (2026) document a 5.5% relative decline in employment of 22–25-year-olds in high-AI-exposure occupations in Sweden by early 2025. By contrast, Kauhanen and Rouvinen (2025, 2026) find no statistically significant overall or seniority-specific effects in Finnish population data. Furthermore, while not testing the age-gradient directly and focusing on the *intensive* rather than the extensive margin, Humlum and Vestergaard (2025) estimate null effects of observed AI-chatbot *adoption* on earnings and hours in a survey-linked subsample of Danish administrative records. Taken together, existing evidence from administrative data suggests that short-run effects are small, heterogeneous across countries, and hard to separate from pre-existing trends.

The Nordic setting is interesting for two reasons. First, adoption was comparatively

quick: by 2025 around 15% of Norwegian firms with ten or more employees used at least one GenAI-related technology, compared with an EU-27 average of 9% (Eurostat, 2026a).[1] Furthermore, in 2025 Norway had the highest population share of GenAI usage (35%) *at work* among European countries, well ahead of the other Nordics (21-27%) and an EU average of merely 15% (Eurostat, 2026b). Therefore, if labor-market effects exist, they should be detectable early on.

Second, Norway is a coordinated labor market whose institutions differ sharply from those of the United States and the United Kingdom and, more subtly, from those of other Nordic countries (Bhuller, Moene, Mogstad, & Vestad, 2022; Iacono, 2018; Mogstad, Salvanes, & Torsvik, 2025). Overall union density in Norway stood at roughly 50% in 2022, and collective-agreement coverage reached approximately 72% of Norwegian employees once the general-application (*allmenngjøring*) mechanism that extends sector-level agreements is taken into account, up from 64% of employees under directly signed agreements (Dølvik et al., 2022; Kjellberg, 2025). Wage setting is further coordinated through the *frontfagsmodell*, a pattern-bargaining system in which trade-exposed manufacturing sectors negotiate first and their settlement serves as a non-binding norm for subsequent settlements across the economy (Dølvik et al., 2022; Government of Norway, 2013). Beyond collective bargaining, Norway has comparatively strict employment protection, including a just-cause requirement, notice periods of one to six months and routine availability of reinstatement. Therefore, labor market adjustment from GenAI is likely to be muted for wages and separations of *incumbent* workers, while new hiring could potentially be more constrained than in more flexible settings. This is in line with existing research: stronger collective bargaining and employment protection tends to cushion incumbents' wages and employment in occupations exposed to automation but accelerate the decline of their occupational employment *shares*, shifting adjustment toward entry rates (Fanfani, 2023; Haapanala, Marx, & Parolin, 2023; Parolin, 2020a, 2020b). This paper provides evidence for both incumbent and overall employment outcomes.

We use the universe of Norwegian employer-employee matched administrative records, linked to occupation-level AI exposure measures, to analyze the early labor-market effects of Generative AI after the introduction of ChatGPT in November 2022. Relative to the existing literature for other countries, we extend the pre-treatment window to January 2015 (roughly doubling the length used in comparable studies), which reveals pre-trend contamination in some specifications and age-bands, and suggests that any apparent post-2022 dynamics in high-AI occupations must be carefully judged as to whether they are mere continuations of pre-existing secular trends. We further complement the canonical within-firm composition design of Brynjolfsson, Chandar, and Chen (2025) and Lodefalk et al. (2026) with two complementary analyses, all run under both the novel Massenkoff and McCrory (2026) and established Eloundou et al. (2024) exposure indices.

We find that at the incumbent level, pooled DiD estimates are subject to noisy pre-trends, and Rambachan and Roth (2023) identified sets include zero for wage, employ-

[1] The usage of general AI technologies is much more widespread, with 29% of Norwegian firms using at least one in 2025. This category however includes pre-GenAI technologies and is not the focus of this paper.

ment and mobility outcomes. To detect hiring-margin adjustments, we deploy three complementary designs. The within-firm composition DiD of the closest related work compares employment in highly-exposed to less-exposed occupations *within* firms. A synthetic difference-in-differences (SDID) compares headcount and wages at the *occupation* level, relaxing the strict parallel-trends assumption and firm-balance restrictions of the within-firm design. A firm-level shift-share approach uses *cross*-firm variation in 2015 occupational portfolios to predict firm-specific AI exposure, and tests whether firms with higher predicted exposure see differential post-treatment employment dynamics. Across all three designs, point estimates for the 22–25 age band are negative but never individually significant. Overall, we observe no significant deterioration of the Norwegian labor market in response to GenAI, neither for early-career workers nor for incumbents. However, point estimates for early-career workers trend more negative late in the post-period, so additional post-treatment data may yet shift the picture.

2 Literature

2.1 Task-based models and evidence from previous automation waves

Modern work on labor-saving technology builds on the task-based framework popularized by Autor et al. (2003), who argue that computers substituted workers in routine cognitive and manual roles but complemented non-routine and complex tasks, leading to structural changes to the composition of jobs in the US economy. Acemoglu and Restrepo (2019) formalize this tension between displacement and reinstatement forces in a general-equilibrium model, showing that automation strictly reduces labor demand unless offset by the creation of new tasks in which labor has a comparative advantage.

A large body of empirical work has examined how these forces played out for previous automation waves. Acemoglu and Restrepo (2022) attribute 50 to 70 percent of observed changes in the US wage structure between 1980 and 2016 to the task displacement of worker groups specialized in routine work in industries experiencing rapid automation. Furthermore, Acemoglu and Restrepo (2020) find negative effects on employment and wages from industrial robot adoption. Other studies have argued for a more nuanced view. Koch, Manuylov, and Smolka (2021) show that on firm level, robot adoption leads to output gains and net job creation within adopters, but substantial job losses at non-adopters. Using patent data, Mann and Püttmann (2023) document a positive influence of automation on employment in local labor markets: while manufacturing employment declined, this was more than compensated by service sector job growth.

2.2 Considerations specific to Generative AI

The theoretical arguments presented for computers and robots largely carry over to any kind of automation technology, including GenAI. The strict distinction between routine and non-routine tasks however blurs, since GenAI is able to perform tasks previously con-

sidered non-routine. This warrants new types of classifications of jobs into those highly at risk from AI and those less so. A growing literature develops such measures of occupational exposure, independently of firm-level adoption of AI technology.

Early measures focused on pre-LLM technologies. Felten, Raj, and Seamans (2021) map advances in AI capabilities to O*NET abilities, and then link these to occupational groups. Pizzinelli, Panton, Mendes Tavares, Cazzaniga, and Li (2023) extend this measure to allow for potential complementarities with labor which could dampen displacement effects. Targeting LLM capabilities specifically, Eloundou et al. (2024) construct an exposure measure by rating whether access to such technologies could reduce the time required to complete a given task by at least 50 percent. Linking tasks to occupations, they estimate that roughly 19 percent of the US workforce could see at least half of their tasks affected, with exposure generally rising in wages.

Moving from predicted exposure to measured use, Handa et al. (2025) map Claude.ai conversations onto O*NET tasks and occupations and document that AI use concentrates in software engineering, writing and analytical tasks, peaks in mid- to high-wage occupations rather than at either extreme of the wage distribution, and splits roughly evenly between automation and augmentation. Massenkoff and McCrory (2026) combine the previous two approaches, and assign higher weights to observed usage in work-related and automated contexts (proxied by whether requests were made via an API). They find that actual usage remains below theoretical capabilities, and observe no increase in unemployment for highly exposed workers since late 2022 in the US; however, they do provide suggestive evidence that hiring of younger workers has slowed down in exposed occupations.

Beyond measurement, Freund and Mann (2026) develop a general-equilibrium model in which occupations bundle tasks and workers sort by comparative advantage across task-specific skills. They project that workers with moderate AI exposure benefit on average while those with high exposure are harmed, with declining returns to analytical skills but rising returns to social ones. However, Acemoglu (2025) argues that based on the current capabilities of AI and given the task-content of contemporary occupations, the macro-economic effects of AI are likely to be modest. Nevertheless, Acemoglu et al. (2026) caution that while AI might not yet be able to substitute for experienced workers, a substantial age-gradient in capabilities exists, which puts entry-level workers in affected occupations at risk of competing with GenAI systems in terms of basic, codified knowledge.

In this study, we primarily focus on the age-gradient prediction, which has received considerable attention in the literature. The wedge between highly and moderately exposed occupations is tested implicitly by considering the latter as part of our control group in most empirical designs. To measure exposure, we rely on the novel index developed by Massenkoff and McCrory (2026), and employ the more widely used measure by Eloundou et al. (2024) as a robustness check.

2.3 Empirical evidence

A first wave of empirical studies focused on controlled field experiments that trace productivity gains from GenAI. Noy and Zhang (2023) show that access to LLMs reduced completion time of mid-level professional writing tasks by 40 percent and raised quality by 18 percent, while compressing productivity differences between workers. Brynjolfsson, Li, and Raymond (2025) find a 15-percent average productivity gain in customer-support jobs, strongly concentrated among less experienced and lower-skilled workers. Similarly, Cui et al. (2025); Gambacorta, Qiu, Shan, and Rees (2024); Peng, Kalliamvakou, Cihon, and Demirer (2023) document strong productivity increases for software development tasks, again with larger gains among less experienced developers.

Naturally, productivity effects might translate into labor demand. A second strand of the literature provides evidence on this with data from surveys, vacancy postings, and online freelancer platforms. Hui et al. (2024) find that on a large online platform the release of ChatGPT, DALL-E 2 and Midjourney reduced both employment and earnings of freelancers in highly affected occupations. Similarly, Teutloff et al. (2025) document 20–50 percent declines in demand for skills such as writing and translation, alongside sharp increases in demand for machine-learning programming and chatbot development. In the U.K., Teeselink (2025) confirms that highly LLM-exposed firms reduce employment, particularly junior hiring in technical and creative roles, and in high-wage segments. However, using a survey of nearly 6,000 senior executives at U.S., U.K., German and Australian firms, Yotzov et al. (2026) document little self-reported impacts on past and future employment.

The latest wave of emerging research uses linked administrative records to identify labor-market effects. While not strictly representative of population-level effects, a widely cited study by Brynjolfsson, Chandar, and Chen (2025) using high-frequency U.S. payroll data documents a 13-percent relative decline in employment of 22–25-year-olds in the most AI-exposed occupations, while employment for less exposed and more experienced workers was stable or growing. In a similar setting, Hosseini and Lichtinger (2025) report a sharp drop in junior employment at firms adopting GenAI, which is driven by slower hiring rather than separations or promotions. A growing body of work is aiming to replicate and extend these results in a European setting using population-wide register data; a literature to which our study aims to contribute to. For Sweden, Lodefalk et al. (2026) combine 4.6 million job advertisements with full-population employer–employee register data, and find an accelerating decline in employment of 22–25-year-olds in high-AI-exposure occupations, reaching 5.5 percent by early 2025 relative to less exposed occupations within the same employers, while employment of workers over 50 rises by 1.3 percent. They interpret their findings as evidence that generative AI shifts hiring composition rather than reducing aggregate demand.

Notably, these results do not replicate in all countries. Using Finnish population data, Kauhanen and Rouvinen (2025) find no statistically significant changes in wages or employment between more- and less-exposed occupations, and Kauhanen and Rouvinen

(2026) find no systematic AI-linked displacement of younger workers. While not testing the age-gradient hypothesis directly, Humlum and Vestergaard (2025) link adoption surveys to Danish administrative labor records and estimate null effects of AI-chatbot adoption on earnings and hours at both worker and workplace levels, although adoption is associated with occupational switching and task restructuring.

Our study contributes to this literature by providing early evidence on the Norwegian population, and extending the analysis to include a long pre-treatment period as well as alternative estimation strategies.

2.4 Institutional moderators

Labor market institutions and related mechanisms have historically shaped the distributional consequences of technological change. Parolin (2020a) shows that in the United States between 1983-2017, higher union coverage cushioned earnings declines in occupations at greater risk of routine-biased automation, but accelerated the decline of their employment shares. Similarly, Parolin (2020b) finds that individual union membership increases the likelihood that workers in routine occupations remain employed and avoid low earnings. Haapanala et al. (2023) argue that stronger organized labor moderates the unemployment effects of robot adoption across countries. Fanfani (2023) uses Italian administrative data matched to industry-wide collective agreements, and finds that contractual wage growth raises actual pay levels but has significantly negative employment effects, consistent with canonical centralized-bargaining models.

Read together, this evidence implies that highly coordinated and protected labor markets such as Norway's should display a muted wage and employment response of incumbents to AI exposure, while adjustments concentrated on the hiring margin might be stronger than in less regulated settings.

3 Data

3.1 Labor market microdata

We use register data provided by Statistics Norway (SSB)[2]. The data is based on the *A-melding* (Statistics Norway, 2026), a report that all Norwegian employers are obliged to file to SSB and the Norwegian tax administration on a monthly basis. It contains income, employment circumstances, and tax-related information. In principle, it covers all employees in any type of work-relationship, with very few exceptions, since January 2015. Our sample covers the period from January 2015 to March 2025; before applying any filters, the worker–month panel comprises approximately 302 million observations covering 3.76 million unique workers in 330,000 firms.

For our analysis, we restrict the sample as follows. First, we only consider information on a worker's *main employment*. The micro-data contains information about all work-

[2]The data were made available by Statistics Norway through an application for access to microdata. Statistics Norway is not responsible for the analyses or the interpretations of the data presented here.

contracts held by a given worker in a given month, including any side jobs or possibly expired contracts or positions that did not accrue any working hours or wages. We disregard such information and hence record at most one job per worker per month, as explained in detail in Annex A.1. Second, we exclude part-time employees in line with Brynjolfsson, Chandar, and Chen (2025); Lodefalk et al. (2026) and restrict the sample to ages 18–66. Third, we exclude rows with missing occupational codes. These are reported on 4-digit level and used to map AI exposures to each worker, as described in the next section. Fourth, we only consider workers in private firms, thus excluding state government, regional or municipal administrations, and state-owned companies.[3] Beyond these overall filters, a number of additional restrictions are later applied flexibly, depending on the specification and as described in Section 4. Some of these restrictions are lifted in subsequent robustness checks.

After restricting to private-sector firms with a full-time main job, the sample contains approximately 113 million worker–months. Imposing additionally that the 4-digit occupational code is non-missing (so that AI exposure can be assigned) and that the worker is aged 18–66 retains a pre-treatment sample (Jan 2015 – Oct 2022) of approximately 1.4 million unique workers. Appendix A.2 reports the filter waterfall in detail and documents balance across AI-exposure quartiles. We use CPI data from Statistics Norway to normalize nominal wages to constant NOK as of September 2025.

3.2 Occupational AI exposure

We do not have access to firm-internal data on AI adoption decisions, nor do we observe proxies such as ICT investments. Therefore, our main analysis relies on the AI exposure index developed by Massenkoff and McCrory (2026), as it combines real-world usage data with higher weights on already automated processes. However, this measure is not yet widely used in the existing literature, so we also provide robustness checks for our main specifications using the index of Eloundou et al. (2024). This allows for more direct comparisons with Brynjolfsson, Chandar, and Chen (2025) and Lodefalk et al. (2026). The measures developed by Felten et al. (2021) and Pizzinelli et al. (2023) were developed before the widespread availability of Generative AI, hence we only use them for selected robustness checks.

We map these indices from US SOC codes to 4-digit occupational codes used in the Norwegian register data, based on crosswalks provided by O*NET and SSB. Across the 396–399 occupations for which exposure can be assigned (depending on the index), the four indices share most of their structure: at the occupational level the rank correlation between Felten et al. (2021), Eloundou et al. (2024) and Massenkoff and McCrory (2026) ranges from 0.79 to 0.89, while Pizzinelli et al. (2023) correlates more weakly (0.44–0.52) with the others. Appendix B reports the full pairwise raw and rank correlations, sum-

[3] This is in line with Brynjolfsson, Chandar, and Chen (2025) and also avoids contamination by a change in Norway's governments at the October 2021 election. Restricting attention to private firms limits the risk that subsequent shifts in public-sector staffing, procurement or funding confound our estimated AI effects.

mary statistics for each index, the distribution of the Massenkoff and McCrory (2026) index across occupations, and the most-exposed STYRK occupations.

4 Empirical strategy

For this study, and in line with the closest related literature, we largely rely on Difference-in-Difference methods, supplemented with Synthetic DiD (Arkhangelsky, Athey, Hirshberg, Imbens, & Wager, 2021) as well as a shift-share design. Our main designs operate at firm and occupation level; a complementary worker-level design is documented in Appendix D.

In the absence of firm-level adoption data for GenAI, and following the related literature (Brynjolfsson, Chandar, & Chen, 2025; Humlum & Vestergaard, 2025; Kauhanen & Rouvinen, 2026; Lodefalk et al., 2026), we treat the public release of ChatGPT in November 2022 as an *availability shock* for Generative AI tools. Treatment intensity is then proxied by occupation-level AI exposure indices from Eloundou et al. (2024) and Massenkoff and McCrory (2026).[4] For the firm- and occupation-level designs presented below, we consider treatment as being in the *highest quartile* of AI exposure. This is in line with similar studies, and aligns with the argument of Freund and Mann (2026) that moderate exposure could in fact be beneficial for workers.

The treatment timing of November 2022 is well-used in the literature and anchored by the first release of ChatGPT. Still, it is important to highlight that beyond text generation, GenAI covers other domains such as image, audio and video creation, and related tools were released as early as January 2021. Similarly, the first generative pre-trained transformer (GPT) was introduced already in June 2018. However, these were early versions with limited output quality and capabilities compared to today, and often they were not widely accessible or easy to use. This is supported by the fact that even by the end of 2023 the use of LLMs in European firms remained at around 2 percent on average, only picking up to 5 and 9 percent in 2024 and 2025 (Eurostat, 2026a). The picture is similar in Norway, where adoption was below 2 percent in both 2021 and 2023 and rose to 10 and 15 percent in 2024 and 2025, respectively. This slow ramp-up supports the treatment timing considered here. Lastly, while Frank, Sabet, Simon, Bana, and Yu (2026) caution that in U.S. unemployment-insurance records labor-market deterioration in AI-exposed occupations already began in early 2022, Lodefalk et al. (2026) in their analysis distinguish effects of a pre-treatment interest-rate hiking cycle from those of GenAI, the latter of which they showed started to materialize only throughout 2024.

Causal identification here relies on several key conditions. Firstly, no anticipation is required for all of the strategies employed in this study, and dictates that subjects must not change their behavior in response to a future treatment. In this regard, Google trends shows a notable 3-fold uptick of online searches related to "Artificial Intelligence" and

[4] In the following, we assume that task contents of occupations in Norway are comparable to those in the US, since the exposure indices used here are based on O*NET tasks and job compositions, and were subsequently mapped to Norwegian occupational codes.

"AI" between November 2022 and mid-2023, while there was relatively little movement before and a strong continuing increase afterwards (Google, 2026). We therefore argue that absent strong public salience prior to November 2022, large-scale anticipatory adjustment by workers (in their educational and career choices) and firms (in hiring practices) is unlikely, although we cannot rule it out from this evidence alone. Testing these assumptions for the worker-side directly is left for future research, as we point out in section 6.

Secondly, in our setting a single event affects all units simultaneously, and identification therefore rests on cross-sectional variation in a baseline moderator (occupation-level AI exposure) interacted with the event timing. Under the *factorial difference-in-differences* framework of Xu, Zhao, and Ding (2024), standard parallel trends (that high- and low-exposure units would have followed the same trajectory absent ChatGPT) identifies only *effect modification*: a statement that the post-event change in outcomes *covaries* with AI exposure. It does not separate the causal effect of AI itself from the effect of any other moderator that happens to be correlated with AI exposure. Occupations in the highest exposure quartile, for instance, are not only high-AI; they are also more concentrated in knowledge work, in larger and more urban firms, and in tech-leaning sectors. Any post-2022 Q4-vs-Q1Q3 divergence will therefore reflect the joint effect of AI *and* any other co-shock or structural moderator that happens to load on the same occupational dimension. The interpretation that AI exposure *causally* moderates the post-2022 response instead requires the stronger *factorial parallel trends* assumption: that the moderator be mean-independent of potential-outcome trends and of other moderators. While this is not directly testable, we extend our pre-trend checks to a period of around twice the length of that used in the related literature, so that confounders with any time-varying signature have more years in which to surface. Furthermore, in Annex C we argue for resilience of our results to concurrent shocks. Beyond that, the four designs used in this study absorb different aspects of moderator-correlated dynamics.

Lastly, the Stable Unit Treatment Value Assumption (SUTVA) is hard to argue for here; if AI displaces workers from highly exposed occupations, they might move to less exposed ones, thereby contaminating the control group. While this attenuates estimated results, we generally consider any results in the following as *general equilibrium effects*. Furthermore, we observe that Norwegian inter-occupation mobility rates over our sample window are low and very stable over time,[5] bounding the magnitude of any such bias.

4.1 Employment outcomes at firm level

To examine whether AI exposure shifts the *composition* of firm-level employment, we follow the within-firm design of Brynjolfsson, Chandar, and Chen (2025) and Lodefalk et al. (2026). The unit of analysis is a firm-quartile-month cell: for each firm f, AI quartile $q \in \{1, 2, 3, 4\}$, and month t in our window, we collapse to the headcount $n_{f,q,t}$ and use $\log(n_{f,q,t}+1)$ as the outcome. The firm panel is restricted to firms observed in every month of the window with an average of at least one worker per quartile per month, so that mean-

[5] Around 70% of Q4 workers stay within Q4 in any given year, with no significant uptick post-treatment.

ingful relative composition can be measured. We estimate the design over two windows: a short window (2019+, six pre-treatment half-years) that emphasizes recent dynamics in the spirit of those studies, and a long window (2015+) that covers the Covid period as well as multiple pre-pandemic years for tighter pre-trend tests. The specification is run both as a pooled version as well as separately across six age sub-bands (22–25, 26–30, 31–34, 35–40, 41–49, 50+), with 22–25 as the focal early-career cohort in line with the related literature. The pooled DiD specification used is:

$$\log(n_{f,q,t}+1) = \alpha_{f,t} + \gamma_{f,q} + \beta_{\text{GPT}}\,\text{Q4}_q \cdot \mathbf{1}[t \geq \text{Nov 2022}] + \varepsilon_{f,q,t}, \tag{1}$$

where Q4_q is an indicator for the high-AI quartile, $\alpha_{f,t}$ a firm-by-month fixed effect, and $\gamma_{f,q}$ a firm-by-quartile fixed effect. The dynamic version replaces the two post-period dummies with a full set of half-yearly interactions:

$$\log(n_{f,q,t}+1) = \alpha_{f,t} + \gamma_{f,q} + \sum_{h \neq 2022\text{H1}} \beta_h\,\text{Q4}_q \cdot \mathbf{1}[t \in h] + \varepsilon_{f,q,t}. \tag{2}$$

The two-way fixed effects play distinct roles. Firm-by-month effects $\alpha_{f,t}$ absorb every shock that hits a firm uniformly across its quartile composition, e.g. firm-specific demand, the oil-revenue boom in petroleum-cluster firms, and aggregate macro-monetary shocks. Firm-by-quartile effects $\gamma_{f,q}$ absorb each firm's persistent baseline composition (some firms are intrinsically Q4-heavy). Identification therefore rests entirely on within-firm changes in the relative size of Q4 versus Q1–Q3 employment. Standard errors are clustered at the firm-quartile level, since treatment varies within firm but is correlated within firm-quartile cells over the months of our window.

The PTA condition here can be interpreted as a flat within-firm Q4-vs-Q1Q3 relative-employment trend in the absence of treatment; we report the joint pre-trend Wald test for $\hat{\beta}_h$, $h < 2022\text{H1}$.

4.2 Synthetic Difference-in-Differences at occupation level

The within-firm design of Section 4.1 isolates a reallocation effect within firms, but conditions on them surviving a balanced 10 year panel on narrow age bands. It cannot speak to occupations whose employment varies through firm entry and exit, nor to AI's gradient on occupation-level wages once compositional churn is netted out. Furthermore, it is still subject to the parallel trends assumption. To complement the analysis, we estimate an analogous treatment effect at the occupation level using the synthetic difference-in-differences (SDID) estimator of Arkhangelsky et al. (2021). For each four-digit occupation o and half-year period p in 2015H1–2024H2,[6] we compute average monthly headcount $n_{o,p}$ and the mean of worker-month log wages $\overline{\log w}_{o,p}$, and run the design separately for each outcome. Treatment is binary at the occupation level: $D_{o,p} = \mathbf{1}[q(o) = 4] \cdot \mathbf{1}[p \geq 2022\text{H2}]$,

[6]The 2025H1 half-year is dropped because our sample ends in March 2025 and would otherwise include only the first three months of the period, possibly biasing headcount due to seasonalities. The post-period therefore contains five half-years, 2022H2 through 2024H2.

with Q1–Q3 occupations as controls. The panel is balanced at the occupation level and restricted to occupations averaging at least 25 workers per pre-treatment period, so that estimated headcount is not driven by small cells. We run the design separately for the pooled cross-age sample and for the same six age sub-bands as in Section 4.1. Inference is via the placebo bootstrap of Arkhangelsky et al. (2021) with 500 replications.

Compared to the within-firm DiD model, SDID on occupational level has two attractive features. First, unit weights are chosen to build a synthetic control whose pre-2022H2 trajectory tracks that of the treated set; the reweighting lets small but high-AI occupations contribute meaningfully to the treated trajectory rather than being dominated by a few large ones. Second, time weights up-weight pre-periods that more closely resemble the post-period, absorbing any linear pre-trend differential between Q4 and Q1–Q3 occupations.

As a falsification exercise, we re-estimate the model with the post-period start moved to 2019H2, three years before ChatGPT and right before the Covid shock, dropping all 2022H2-onwards observations. A non-zero placebo ATT in this window would indicate that the synthetic-control matching is not adequately absorbing pre-trend differentials.

4.3 Shift-share exposure at the firm level

The within-firm and SDID designs both identify off variation in Q4 versus Q1–Q3 employment, either within firms or across the cross-section of occupations. A third, complementary design uses variation across *firms* in their *pre-period* occupational portfolios: firms whose 2015 workforce was concentrated in occupations that turned out to be highly AI-exposed should, if AI displacement operates at the firm level, see total employment of their workforce decline relative to firms who were less exposed. By using across-firm variation in pre-existing portfolios rather than within-firm reallocation or within-occupation comparisons, this shift-share design provides a complementary angle to the two designs above.

We estimate the design separately for each age band a. For each firm f and age band, predicted exposure is

$$z_{f,a} = \sum_{o} s^{2015}_{f,a,o} \cdot A_o, \tag{3}$$

where $s^{2015}_{f,a,o}$ is firm f's 2015 share of its age-a workforce employed in occupation o, and A_o is the occupation-level AI score from the chosen index. Both the share matrix *and* the outcome below are restricted to age band a: the design tests whether 2023+ age-a employment in firm f declines in proportion to its age-a-specific predicted exposure.[7] We standardize $z_{f,a}$ to mean zero and unit variance within each age band; β_a below therefore reports the effect of a one-standard-deviation increase in firm-level predicted exposure.

The outcome panel is at firm–year frequency: $n_{f,a,t}$ is the headcount of age-a workers

[7] We acknowledge that a is subject to drift via aging. However, age is typically correlated with job *seniority* and hence task content. Thus, this design still captures a relevant margin: for example, firms with many early-career workers highly exposed to AI should see a decline in their overall employment if AI reduces demand for the tasks typically performed by such workers.

in firm f in year t, and we model $\log(n_{f,a,t}+1)$ over 2015–2024. The firm panel is balanced at the firm level over the ten-year window and restricted to firms with at least five age-a workers in 2015, so that the share matrix in equation (3) is computed on a non-trivial baseline.

The pooled specification is

$$\log(n_{f,a,t}+1) = \alpha_{f,a} + \delta_{a,t} + \beta_a \, z_{f,a} \cdot \mathbf{1}[t \geq 2023] + \varepsilon_{f,a,t}, \tag{4}$$

with firm fixed effects $\alpha_{f,a}$ and year fixed effects $\delta_{a,t}$, both estimated within each age band. The dynamic version replaces the post-treatment indicator with a full set of year interactions:

$$\log(n_{f,a,t}+1) = \alpha_{f,a} + \delta_{a,t} + \sum_{k \neq 2022} \beta_{a,k} \, z_{f,a} \cdot \mathbf{1}[t = k] + \varepsilon_{f,a,t}, \tag{5}$$

with year 2022 as the omitted reference. Standard errors are clustered at the firm level.

Identification rests on a cross-firm parallel-trends assumption: firms with different 2015 occupational portfolios would have followed parallel trajectories in $\log(n_{f,a,t}+1)$ absent the post-2022 AI shock. The corresponding identifying threat is that pre-existing secular trends (skill-biased technology adoption, sectoral reallocation, scale-up of high-exposure industries) differentially favored firms with high-exposure 2015 portfolios for reasons unrelated to the November 2022 ChatGPT shock. While this is inherently untestable, we provide joint Wald tests on the pre-treatment event-study coefficients $\beta_{a,k}$ over a long pre-treatment period.

5 Results

5.1 Descriptive patterns of employment by AI exposure

We begin by examining the cross-sectional composition of the four quartiles of AI exposure across the pre-period, summarized in Appendix Tables 5 and 6. This reveals systematic differences in terms of observed worker-characteristics. First, the most-exposed quartile (Q4) is the largest in worker–months (29.0M, against 19.5M in Q1), indicating that AI-exposed occupations cover a substantial share of Norway's private-sector workforce rather than a niche subset. Second, Q4 workers have a mean age of 41.5 (the second-oldest among the four bins) likely reflecting an overrepresentation of occupations that require tertiary education; yet Q3 is older still by three to four years and earns substantially more, consistent with a concentration of senior knowledge-work occupations. Third, Q4 has the lowest immigrant share at 12%, mirroring immigrants' greater representation in less-exposed manual jobs.

We next document how relative employment across AI exposure quartiles has evolved around ChatGPT's release. To inspect the dynamic patterns most directly relevant to the recent literature (Brynjolfsson, Chandar, & Chen, 2025; Lodefalk et al., 2026), we plot indexed monthly employment (Oct 2022 = 100) by AI exposure quartile within each age band

in Figure 1. Two features stand out. First, in the youngest age band (22–25), employment in the highest-exposure quartile (Q4) declines markedly from late 2023 onward while the three lower quartiles remain broadly stable. For the 26–30 and 41–49 bands, Q3 and Q4 share a common upward trend through end 2022 but Q4 reverses post-treatment while Q3 keeps growing. Second, the same Q4-relative-decline pattern is largely absent or even reversed for the remaining age bands.

These descriptive patterns indicate a possible age gradient in the employment response to Generative AI, which motivates the formal analyses that follow. The next three sections estimate complementary difference-in-differences designs, each run separately by age band: the within-firm composition design (Section 5.3) tests for compositional shifts within balanced firm panels; the synthetic difference-in-differences (Section 5.4) moves the analysis to the occupation level, sidestepping the sample shrinkage that the within-firm design's firm-balance restriction imposes on narrow age bands; and the shift-share design (Section 5.5) exploits cross-firm variation in pre-treatment occupational portfolios. A worker-level event study (summarized in Section 5.2 and reported in full in Appendix D) serves as a complementary check on incumbent outcomes, the margin where Norway's institutional setting gives the strongest priors of no effect.

5.2 Worker-level analysis

We hypothesize that the Norwegian institutional context (high union density, strict employment protection, and collectively-bargained wages) buffers *incumbent* workers from a direct AI displacement effect, so that any post-2022 response should concentrate at the firm and occupation margins examined below rather than at the worker level. As a complementary check, we run a worker-level pooled DiD and event study on a 10-year balanced panel of full-time private-sector employees, with four outcomes spanning: log monthly wage, months employed in the year, and binary indicators for changing firm or changing occupation. Three of the four pooled coefficients are significant at the 1% level, but event-study pre-trends are pronounced and the HonestDiD bounds of Rambachan and Roth (2023), which adjust the post-treatment effect for the magnitude of the observed pre-period drift, include zero for all four outcomes. Furthermore, the pooled coefficients are unstable across robustness specifications. We therefore cannot reject a zero effect of AI exposure on incumbent workers, in line with our institutional priors. Given this noise and the broader literature's focus on employment numbers, we report the full specification, results, and robustness only as Appendix D.

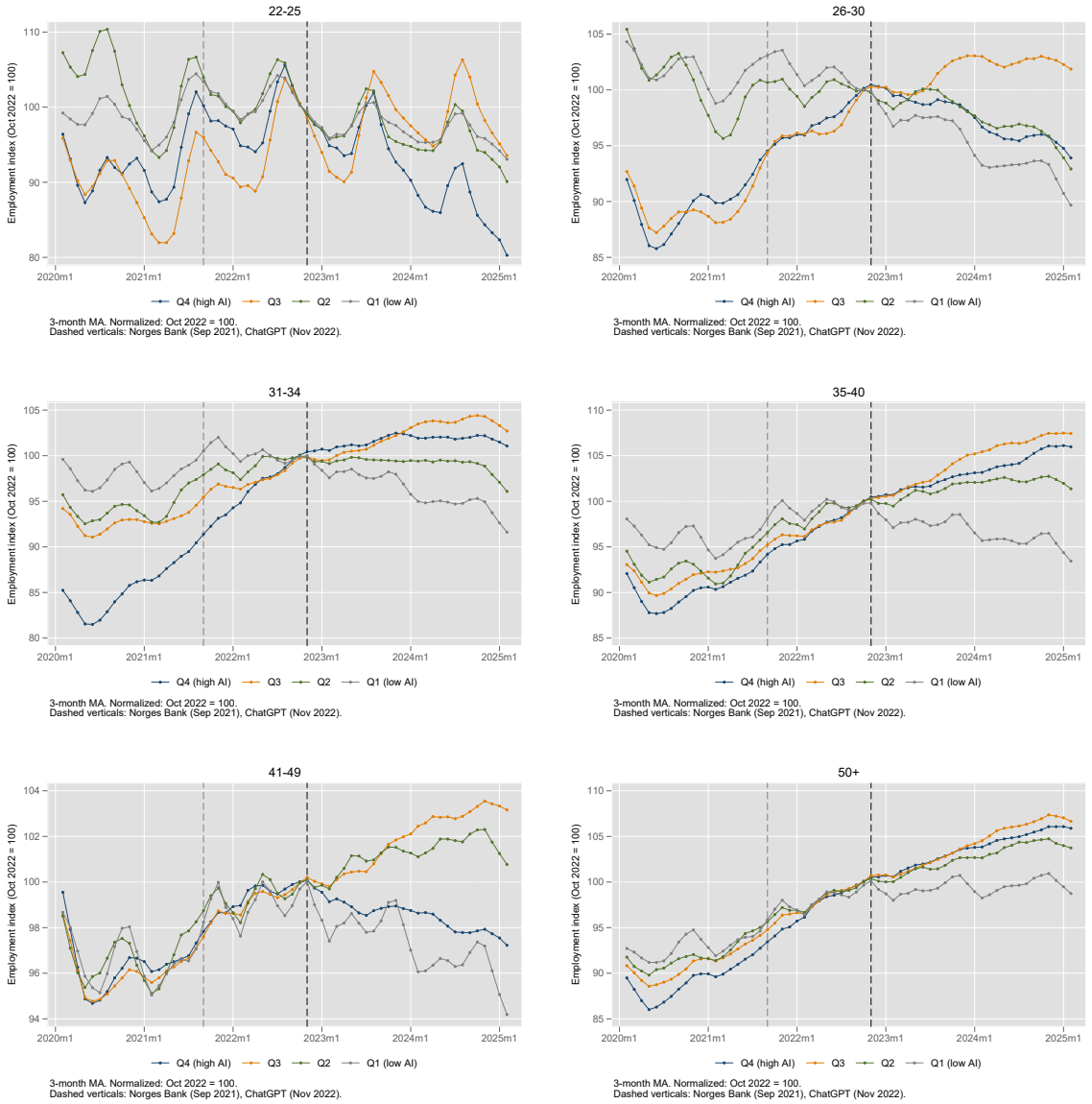


Figure 1: Indexed monthly employment within AI exposure quartiles, by age band (Oct 2022 = 100). Each panel plots employment trajectories for the four quartiles within a single age band; vertical reference at November 2022 (ChatGPT release). Massenkoff and McCrory (2026) exposure index, contemporaneous occupation. Sample: private sector, full-time main job, ages 18–66, non-missing 4-digit occupation.

5.3 Within-firm composition

Figure 2 contrasts Q1 versus Q4 employment trajectories separately for workers of ages 22-25 and 26-plus workers, indexing each series to October 2022. The young-worker gap between Q4 and Q1 widens markedly post-ChatGPT: by early 2025, employment in high-AI occupations is roughly twenty percentage points lower (relative to October 2022) than employment in low-AI occupations among young workers, while the analogous gap among older workers is of opposite sign. This pattern is consistent with the literature's age-gradient prediction (Brynjolfsson, Chandar, & Chen, 2025; Lodefalk et al., 2026).

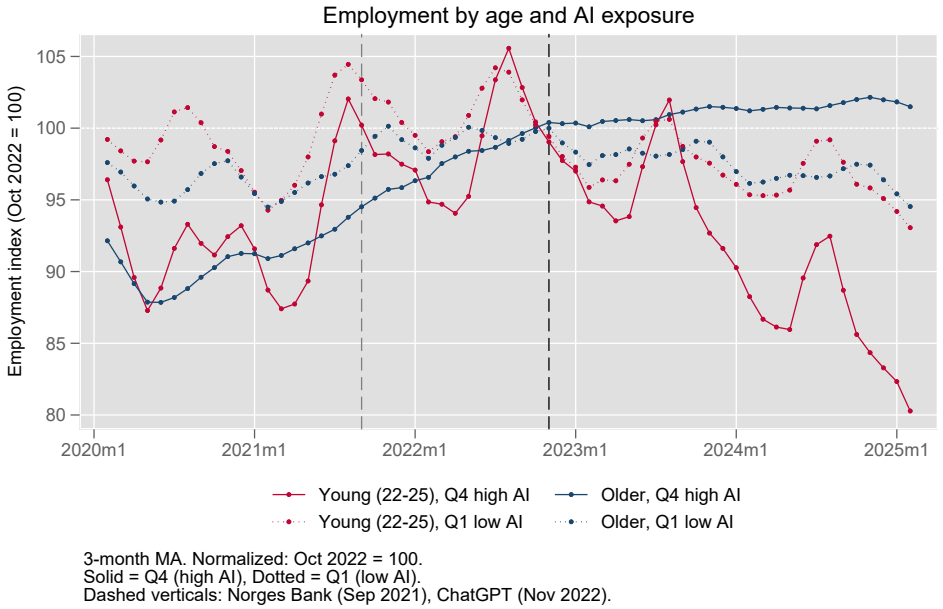


Figure 2: Indexed monthly employment in Q4 (high) and Q1 (low) AI-exposure quartiles, separately for workers between 22-25 years and older workers, normalized to October 2022 = 100. Massenkoff and McCrory (2026) exposure index, contemporaneous occupation. Sample: private sector, full-time main job, ages 18–66, non-missing 4-digit STYRK occupation.

To take this descriptive pattern to a formal test, we estimate the within-firm composition design of Section 4.1. Recall that the unit of analysis is a firm × AI quartile × month cell with outcome $\log(n_{f,q,t} + 1)$, and the panel is balanced over the 2015–2024 window retaining only firms with at least one worker per quartile (Q4 and Q1-Q3) per month. Identification thus relies on within-firm reallocation across exposure quartiles. The firm panel that survives the balanced-panel and quartile-presence filters contains 6,377 firms in the all-ages cross-section and shrinks for narrow age bins, reaching only 45 firms in the 22–25 panel. Table 1 reports the results for equation (1) for each age band, together with the joint pre-trend F-statistic from the corresponding event study.

Table 1: Pooled within-firm DiD coefficients by age band on the long (2015+) window (Massenkoff and McCrory (2026) index).

Age	$\hat{\beta}_{\text{GPT}}$	SE	p-value	Pre-trend $F(14)$	N	Firms
All ages	0.0158^{***}	0.0047	0.001	10.32^{***}	3,137,484	6,377
22–25	-0.0704	0.0656	0.284	0.41	22,140	45
26–30	0.0717^{**}	0.0284	0.012	1.46	139,236	283
31–34	0.0453^{*}	0.0274	0.099	1.50	153,504	312
35–40	0.0144	0.0179	0.420	2.22^{***}	331,608	674
41–49	-0.0228^{**}	0.0115	0.048	2.48^{***}	713,892	1,451
50+	0.0695^{***}	0.0075	0.000	8.68^{***}	1,260,504	2,562

Each row is a separate within-firm DiD (equation (1)) restricted to firm-quartile-month cells in the named age band. The pre-trend F-statistic is the joint Wald test on the dynamic-specification half-yearly coefficients $\hat{\beta}_h$ for $h <$ 2022H1 in equation (2). SE clustered at firm × quartile. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The ChatGPT coefficient $\hat{\beta}_{\text{GPT}}$ is negative only for the 22–25 and the 41–49 age bands, and only significant for the latter. For the 22–25 panel, the pre-trend test passes cleanly ($F(14) = 0.41$), but the estimated effect of a 7 log-point relative drop in within-firm Q4 employment of 22–25-year-olds following ChatGPT is *not* statistically significant from zero, and estimated on only a very small number of identifying firms.

Furthermore, the 50+ panel returns a positive and statistically significant ChatGPT coefficient ($+0.069^{***}$). On its own this would suggest that older workers are concentrated into high-AI occupations after late 2022. The strong pre-trend ($F(14) = 8.68$) makes this a continuation of a long secular shift rather than a 2022 break, however. The same observation holds for the coefficient of the pooled all-ages panel.

Figure 3 plots the half-yearly event-study coefficients of equation (2) for three representative age bands: pooled (all ages), 22–25 and 50+. The pooled panel shows a steady upward pre-trend that the high F flags: high-AI shares within firms have been climbing since 2015, albeit with a pronounced slowdown after November 2022. The 22–25 panel is by contrast remarkably flat through 2022 and drops below zero post-treatment; coefficients are however estimated noisily and confidence intervals are very wide, reflecting a lack of power due to the severe sample-shrinkage in this narrow age band. Finally, the 50+ panel shows a smooth, near-linear continuation of its pre-2022 trajectory into the post-treatment period, confirming that the pooled coefficient in this age-group cannot be read as a causal effect of GenAI.

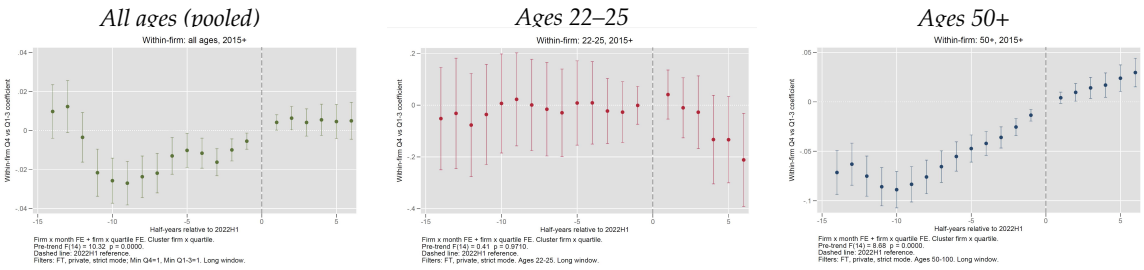


Figure 3: Within-firm composition event-study coefficients $\hat{\beta}_h$ from equation (2), three age bands. Reference period 2022H1. Sample: balanced firm panel (2015–2024) with at least 1 worker per AI quartile per month; Massenkoff and McCrory (2026) index. 95% confidence intervals on standard errors clustered at firm $\times$ quartile.

Appendix E reports the same age-band pooled DiD on the alternative shorter (2019+) window and with an additional dummy for the interest rate hike in 2021 following Lodefalk et al. (2026), the analogous age-band decomposition under the Eloundou et al. (2024) index, and the four cohort-by-cohort event studies for the 26–30 through 41–49 bands omitted from the main text. The 22–25 ChatGPT coefficient under the shorter window has the same negative sign but remains insignificant, under the Eloundou index the same coefficient is also negative and of similar magnitude (-0.066), but remains insignificant. The cross-spec evidence on within-firm 22–25 displacement is therefore directionally consistent but not conclusively identified in either specification. The additional event-studies largely exhibit similar pre-trend issues as the all-age and 50+ panels.

A natural concern here is that the strict cross-quartile filter, which mirrors Brynjolfsson, Chandar, and Chen (2025) but in a much smaller population, drives the small identifying sample in narrow age bands. However, closing the precision gap relative to Lodefalk et al. (2026) is harder than it appears. First, Sweden's workforce is roughly twice the size of Norway's and Lodefalk et al. (2026) additionally include public-sector employees and part-time workers, so their cell count is several times larger than ours independent of any filter choice. Second, relaxing our filter to admit firms with partial cross-quartile presence (closer to their approach) inflates the panel substantially, but the additional firms typically have little to no within-firm Q4-vs-Q1–Q3 turnover and therefore contribute little identifying variation. Under OLS on $\log(n+1)$ the predominantly-zero cells tend to introduce mechanically more noise, without delivering tighter confidence-intervals. We therefore retain the stricter filter as our headline specification, and leave a Poisson PML variant, which handles zero-heavy panels more cleanly and could enable a meaningfully looser filter, to a future iteration of the paper.

To sum up, under the current specification we find marginal evidence for the hypothesis that GenAI has contributed to a decline of employment of young workers in Q4 occupations. However, results are not significant and very noisy. In any case, the within-firm design conditions on firms surviving a balanced 10-year panel with substantial cross-quartile presence; these are large, structurally diversified firms whose composition margin is well-measured but who may not be representative of the Norwegian private sector. The next section relaxes this restriction by moving the design to the occupation level.

5.4 Synthetic difference-in-differences at occupation level

Section 4.2 introduced the occupation-level Synthetic DiD design, which complements the firm-level analysis by dispensing with the within-firm balance that drove the firm sample down to 45 cells in the 22–25 cohort. This also opens up the possibility of detecting effects driven by firm entry or exits. The base panel covers 396 occupations spanning 2015H1 to 2024H2 (Q1: 100, Q2: 98, Q3: 99, Q4: 99). We impose the requirement that each occupation averages at least 25 workers per pre-treatment half-year to avoid effects from small and noisy job titles;[8] this retains 89 treated (Q4) and 261 control (Q1–Q3) occupations in the all-ages run and narrows to 46 treated and 142 control occupations in the 22–25 panel. For each cell we compute monthly headcount and the mean of worker-month log monthly cash wages, and we estimate a Synthetic DiD once for each outcome and each of the seven age bands. As a falsification exercise, we re-estimate all specifications with the post-period start moved to 2019H2 (dropping all post-2022H2 data), three years before ChatGPT and immediately before the Covid shock: a non-zero ATT in the 2019H2 backdating would indicate that the synthetic-control matching is not adequately absorbing pre-trend differentials.

Table 2 reports the SDID ATT estimates and standard errors from the placebo bootstrap of Arkhangelsky et al. (2021) (500 replications). Of the fourteen real-treatment cells (seven

[8] This restriction will be relaxed as a robustness check in a future iteration of this paper.

age bands × two outcomes), two reach $p \leq 0.05$: the 31–34 cell on headcount ($\hat{\tau}_n = +0.039$) and the 50+ wage cell ($\hat{\tau}_w = +0.014$). The remaining twelve cells are statistically indistinguishable from zero. The cross-design comparison for the youngest cohort is informative: the SDID 22–25 headcount ATT of -0.013 shares the negative sign with the within-firm -0.07 from Section 5.3 but lies an order of magnitude smaller in absolute value, with a 95% confidence interval that comfortably includes zero ($[-0.083, +0.058]$). The 2019H2 backdating produces a 22–25 headcount ATT of -0.030 (larger in magnitude than the real-treatment ATT) which is itself statistically indistinguishable from zero but rules out interpreting the post-2022 estimate as an AI effect. We conclude that, at the occupation level, the within-firm 22–25 result does not have an analog of comparable strength.

Table 2: Synthetic difference-in-differences ATTs at the occupation level, by age band and outcome (Massenkoff and McCrory (2026) index). Real-treatment specifications use post-2022H2; 2019H2 backdating drops post-2022H2 observations and treats 2019H2 as the false treatment cutoff.

	Headcount: $\log(\bar{n}+1)$				Wage: $\overline{\log w}$			
Age band	Real	(SE)	2019H2 backdating	(SE)	Real	(SE)	2019H2 backdating	(SE)
All ages	-0.001	(0.013)	-0.001	(0.023)	0.017	(0.012)	0.016	(0.012)
22–25	-0.013	(0.036)	-0.030	(0.040)	-0.005	(0.016)	-0.009	(0.016)
26–30	0.015	(0.021)	0.030	(0.027)	0.013	(0.010)	0.018^{*}	(0.011)
31–34	0.039^{**}	(0.018)	0.042^{*}	(0.022)	0.018^{*}	(0.009)	0.012^{**}	(0.006)
35–40	0.013	(0.015)	0.026	(0.019)	0.007	(0.008)	0.014^{**}	(0.005)
41–49	-0.022	(0.015)	0.010	(0.016)	0.009	(0.010)	0.014^{**}	(0.006)
50+	0.011	(0.014)	0.030^{*}	(0.016)	0.014^{**}	(0.007)	0.009^{*}	(0.005)

Each cell reports an ATT estimate from a separate Arkhangelsky et al. (2021) SDID run on the corresponding occupation × half-year panel. Standard errors via placebo bootstrap (500 replications). ** $p < 0.05$, * $p < 0.10$.

The two unexpected positives (31–34 headcount and wage) are not predicted by any reading of the literature and we treat them with caution: magnitudes are small (4 log points on headcount, 2 log points on wage) and neither replicates under the Eloundou et al. (2024) index (Appendix Table 15: $+0.020$ headcount with $p = 0.29$; $+0.012$ wage with $p = 0.21$). The 50+ wage point estimate is similarly small and similarly fails to replicate under Eloundou ($+0.008$, $p = 0.24$). We therefore do not view these isolated positives as evidence of an AI productivity premium for mid-career or older workers.

Figure 4 plots the synthetic-control trajectory for the two age bands of main interest (22–25 and 50+) and for both outcomes. The 22–25 panels show a treated trajectory that tracks the synthetic control closely throughout the pre-period and exhibits no visible break at 2022H2 on either outcome. The 50+ panels likewise show treated and synthetic series that move together through the post-period; the small positive ATT on 50+ wage emerges from a gradual divergence after 2022H2, which however is of comparable magnitude to the visible noise in the synthetic-control match. Note also that the SDID time weights concentrate on the most recent pre-treatment observations, yet the synthetic control tracks the treated series across the full 2015H1–2022H1 pre-period. Therefore, unit weights deliver a structurally credible match over a substantially longer horizon than the time-weighting alone would require.

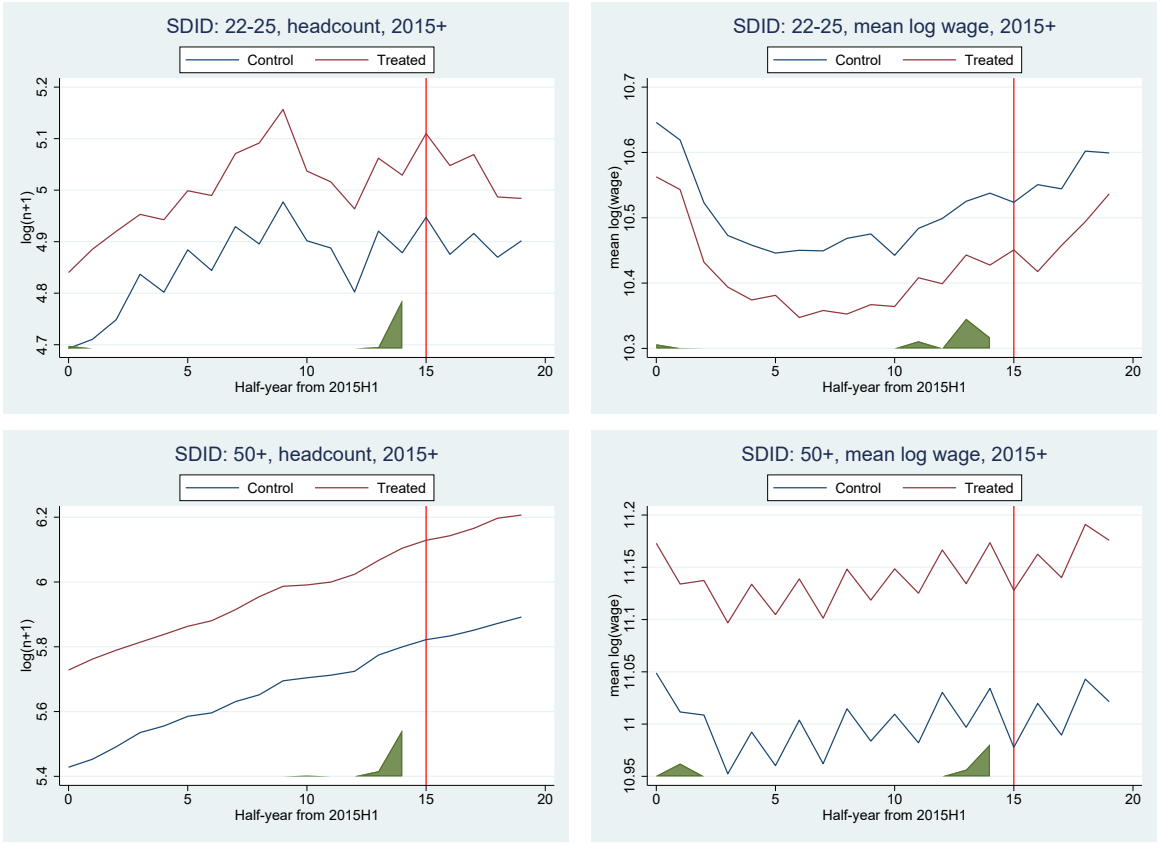


Figure 4: Synthetic-control trajectories for the 22–25 and 50+ age bands, both outcomes. Vertical reference at 2022H2. The green shaded area represents estimated time-weights. Massenkoff and McCrory (2026) index.

Appendix F reports the same design under the Eloundou et al. (2024) index (with even weaker results), the trajectories for the five age bands omitted here, as well unit-weight diagnostics for the synthetic control (showing that control weights are roughly equally distributed across Q1 to Q3 of AI exposure). It also reports two further robustness exercises: redefining treatment as above-median (rather than top-quartile) AI exposure, and re-estimating the design on a publicly available SSB occupation panel that extends through 2026Q1 but with lower age-granularity. Both yield real-treatment headcount ATTs that are systematically dominated by their 2019H2 backdating placebos, reinforcing the reading that post-2022 dynamics in high-AI occupations are largely a continuation of pre-existing trends.

5.5 Shift-share at the firm level

Section 4.3 introduced the shift-share design at the firm level, which uses cross-firm variation in 2015 occupational portfolios to predict post-2022 employment dynamics. Unlike the within-firm design of Section 5.3, identification here does not rely on relative composition *within* the same firm, and unlike the SDID of Section 5.4, it operates at the firm rather than the occupation level. The design therefore complements the previous two with a somewhat different angle that keeps a firm-specific perspective and is not contaminated

by firm entry or exit but still allows a somewhat larger sample due to less restrictive filters than the within-firm design. The filters imposed here are balanced presence of firms over 2015–2024 and a 2015 baseline of at least five age-a workers per firm. For the 22–25 cohort this raises the identifying firm count from 45 (within-firm) to 82.

Table 3 reports the pooled DiD coefficient $\hat{\beta}_a$ from equation (4) by age band, under both AI indices, together with standard errors, the joint pre-trend Wald F-statistic, the associated p-value, and firm counts.

Table 3: Shift-share pooled DiD coefficients by age band, both AI indices.

	Massenkoff and McCrory (2026)				Eloundou et al. (2024)				
Age band	$\hat{\beta}_a$	SE	$F(7)$	p_{pre}	$\hat{\beta}_a$	SE	$F(7)$	p_{pre}	Firms
All ages	0.006	0.010	4.60^{***}	< 0.001	0.002	0.011	2.69^{***}	0.009	3,381
22–25	-0.101	0.069	1.31	0.254	-0.080	0.069	1.02	0.424	82
26–30	0.027	0.037	2.10^{**}	0.044	0.056	0.044	1.08	0.376	257
31–34	0.130^{***}	0.043	2.50^{**}	0.017	0.089	0.064	1.61	0.133	204
35–40	-0.022	0.030	1.61	0.129	-0.045	0.039	1.39	0.208	413
41–49	0.022	0.024	0.94	0.473	0.018	0.030	0.52	0.821	777
50+	0.049^{**}	0.023	1.46	0.180	0.037	0.027	1.03	0.405	859

$\hat{\beta}_a$ is the pooled-DiD coefficient from equation (4). $F(7)$ is the joint Wald test that all seven pre-treatment interactions equal zero, with stars denoting rejection of parallel trends; p_{pre} is the corresponding p-value. Firm counts are identical across indices because firm selection uses index-independent filters. Standard errors clustered at the firm level. $^{***}\,p < 0.01$, $^{**}\,p < 0.05$, $^{*}\,p < 0.10$.

The 22–25 row delivers the design's most interpretable estimate: under Massenkoff and McCrory (2026), $\hat{\beta}_{22\text{–}25} = -0.101$ on a clean pre-trend ($p_{\text{pre}} = 0.25$), with the pooled effect not significant ($p \approx 0.14$) on a panel of 82 firms. The negative sign is directionally consistent with the within-firm -0.070 and the SDID -0.013 obtained in Sections 5.3 and 5.4; on the larger shift-share sample the per-SD magnitude is ten log points (although magnitudes are not directly comparable), though confidence intervals remain wide. Under Eloundou et al. (2024) the same coefficient is -0.080 with an equally clean pre-trend, again not significant. Figure 5 (top row) shows the event-study trajectories under both indices: coefficients are flat through 2022, drift modestly negative in 2023, and the 2024 coefficient is the only post-treatment estimate to fall significantly below zero. This late-period acceleration mirrors the pattern in the within-firm 22–25 event study in Figure 3 and observations from Brynjolfsson, Chandar, and Chen (2025); Lodefalk et al. (2026).

The 50+ row is methodologically more credible than the within-firm 50+ result: Under Massenkoff and McCrory (2026), $\hat{\beta}_{50+} = +0.049$ ($p \approx 0.035$) on a panel of 859 firms with a clean pre-trend ($p_{\text{pre}} = 0.18$). The corresponding within-firm coefficient of $+0.070^{***}$ rests on a dirty pre-trend ($F(14) = 8.68$, $p < 0.001$), so the within-firm reading there could not be cleanly attributed to a 2022 break. The shift-share estimate has neither this problem nor an unstable pre-period: Figure 5 (bottom row) shows the trajectory drifting modestly upward through 2023 and 2024, with only the last two coefficients individually marginally significant. Under Eloundou et al. (2024) the sign is preserved ($+0.037$) and the pre-trend remains clean, although statistical significance is lost.

The remaining age bands offer less interpretable results, and we treat the 31–34 bin as

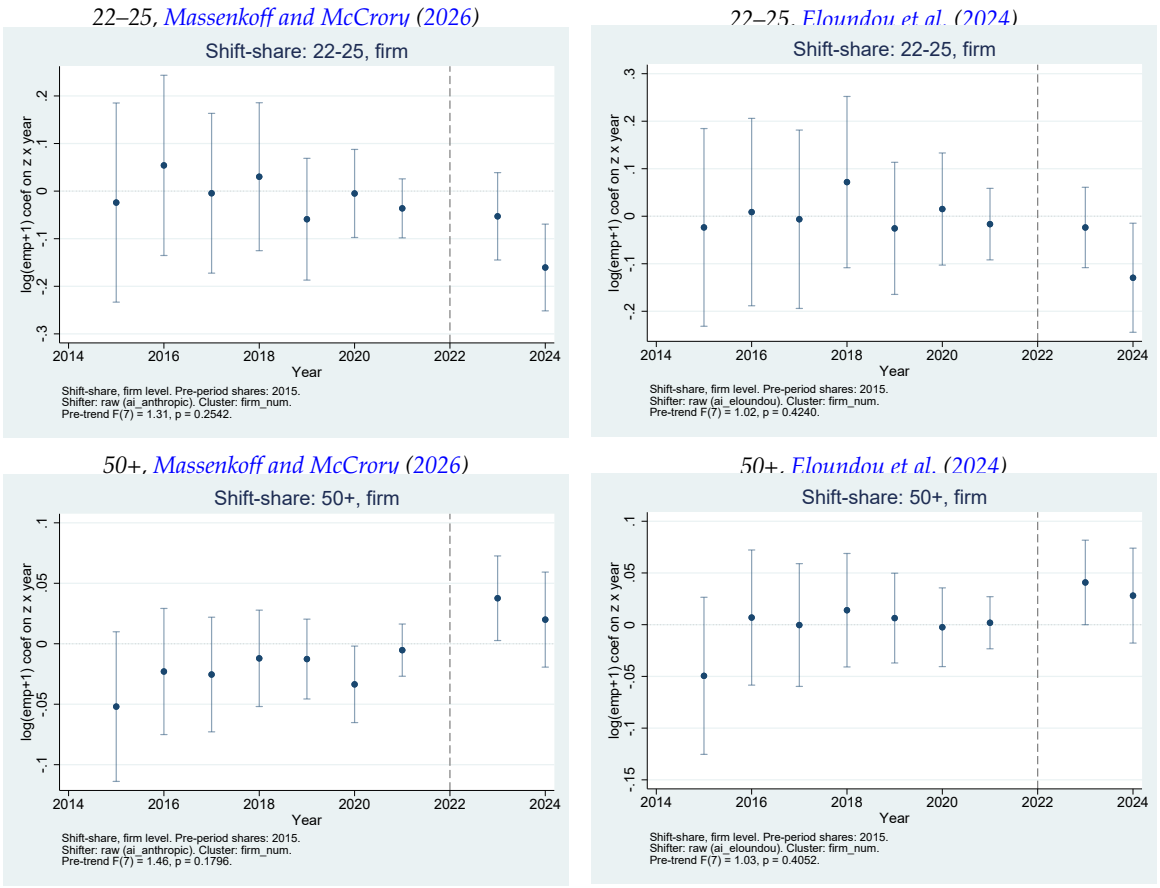


Figure 5: Shift-share event-study coefficients $\hat{\beta}_{a,k}$ from equation (5), for the two age bands of interest under both AI indices. Reference year 2022. 95% confidence intervals based on standard errors clustered at the firm level.

fragile, in line with the same caveat that applied to the SDID 31–34 cells in Section 5.4. Appendix G reports the event-study charts for the five age bands omitted from Figure 5.

6 Conclusion

6.1 Summary

Our four difference-in-differences designs deliver a cautious reading of the early labor-market consequences of Generative AI. At the worker level, a balanced ten-year panel of incumbents and four outcomes (log wage, months employed, firm- and occupation-mobility) yields DiD estimates subject to non-parallel pre-trends. Robust identified sets following Rambachan and Roth (2023) include zero across all four outcomes, implying that post-treatment estimates cannot be distinguished from zero (Appendix D). We interpret this as a null effect on incumbent labor-market outcomes, in line with our institutional priors and the null worker-level effects of Humlum and Vestergaard (2025) in Danish administrative data.

To detect adjustments at the hiring margin, we first replicate the within-firm composition design of Brynjolfsson, Chandar, and Chen (2025), which contrasts the relative size

of high- and low-exposure quartiles within the same firm across several age bands. For workers aged 22–25, the post-ChatGPT estimate is negative but not statistically significant, and identified on a small panel of only 45 firms that survive our filters and balance requirements. The point estimate is directionally consistent with the negative findings of Brynjolfsson, Chandar, and Chen (2025) for the United States and Lodefalk et al. (2026) for Sweden; however our confidence interval is too wide to give a sharp estimate for Norway. Nevertheless, the 22–25 event-study trajectory shows a clean pre-trend and accelerates downwards towards the end of the sample period. The remaining age bands are estimated on larger firm samples but either deliver dirty pre-trends, null point estimates, or both: the all-ages and 50+ panels in particular show steady secular increases in high-AI employment that our longer pre-period reveals.

Because the within-firm design conditions on firms surviving a strict balanced ten-year panel with sufficient cross-quartile presence, we complement it with a synthetic difference-in-differences at the occupation level, which covers a substantially larger population of workers but admits effects from firm-entry and -exit. The 22–25 headcount ATT is again negative but not significant, while a backdating placebo yields a larger absolute estimate. Extending the backdating exercise to all seven age bands further reveals that placebo ATTs match or exceed the real-treatment estimates. This indicates that the apparent post-2022 dynamics in high-AI occupations are, at least in part, continuations of pre-existing secular trends rather than clean AI-period breaks.

A third design, a firm-level shift-share, asks whether age-cohort employment evolved differently in firms whose 2015 workforce was concentrated in highly exposed occupations, identifying off cross-firm variation rather than within-firm reallocation or cross-occupation contrasts. For the 22–25 cohort the estimate is again negative on a clean pre-trend but not statistically significant, although identified on a larger panel of 82 firms. For workers aged 50 and over, a one-standard-deviation increase in firm-level predicted exposure is associated with a roughly 5% post-ChatGPT increase in employment ($p \approx 0.035$) on a clean pre-trend. This points in the same direction as the within-firm 50+ result, and is consistent with evidence from related studies.

Taken together, this paper provides population-level evidence on the early labor-market effects of Generative AI in Norway, and complements the within-firm composition analyses of the closest related work with two further designs, all run under both the novel Massenkoff and McCrory (2026) and established Eloundou et al. (2024) exposure indices. Across all three firm- and occupation-level designs, we find directionally consistent but non-robust evidence of a deterioration of early-career employment: 22–25 point estimates are negative but never individually significant, thus consistent only in direction with the negative findings of Brynjolfsson, Chandar, and Chen (2025) and Lodefalk et al. (2026). Furthermore, we extend the pre-treatment window to January 2015, roughly doubling the length used in comparable studies, which reveals pre-trend contamination in some specifications and cohorts. This indicates that longer pre-treatment windows combined with multiple complementary identification strategies appear essential for distinguishing

AI-period breaks from secular trends in administrative data. Overall, we observe no significant deterioration of the Norwegian labor market in response to GenAI, neither for early-career workers nor for incumbents. However, accelerating event-study coefficients towards the end of our sample suggest that it may simply be too early to detect any robust effects.

References

Acemoglu, D. (2025). The simple macroeconomics of ai. *Economic Policy*, *40*(121), 13–58.

Acemoglu, D., Autor, D., & Johnson, S. (2026). Building pro-worker artificial intelligence. *NBER Working Paper 34854*.

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3–30. doi: 10.1257/jep.33.2.3

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, *128*(6), 2188–2244. doi: 10.1086/705716

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. *Econometrica*, *90*(5), 1973–2016. doi: 10.3982/ECTA19815

Alstadsæter, A., Bratsberg, B., Eielsen, G., Kopczuk, W., Markussen, S., Raaum, O., & Røed, K. (2020). *The first weeks of the coronavirus crisis: Who got hit, when and why? Evidence from Norway* (NBER Working Paper No. 27131). National Bureau of Economic Research. doi: 10.3386/w27131

Alstadsæter, A., Bratsberg, B., Markussen, S., Raaum, O., & Røed, K. (2025). Social gradients in employment during and after the COVID-19 pandemic. *Journal of Economic Inequality*. doi: 10.1007/s10888-024-09645-6

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. doi: 10.1257/aer.20190159

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. doi: 10.1162/003355303322552801

Bhuller, M., Moene, K. O., Mogstad, M., & Vestad, O. L. (2022, November). Facts and fantasies about wage setting and collective bargaining. *Journal of Economic Perspectives*, *36*(4), 29–52.

Brynjolfsson, E., Chandar, B., & Chen, R. (2025, November). *Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence* (Stanford Digital Economy Working Paper). Stanford Digital Economy Lab.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, *140*(2), 889–942. doi: 10.1093/qje/qjae044

Cui, K. Z., Demirer, M., Jaffe, S., Musolff, L., Peng, S., & Salz, T. (2025). The effects of generative AI on high-skilled work: Evidence from three field experiments with software developers. *Management Science*. doi: 10.1287/mnsc.2025.00535

Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, *189*, 104235. doi: 10.1016/j.jpubeco.2020.104235

Dølvik, J. E., Marginson, P., Alsos, K., Arnholtz, J., Meardi, G., Müller, T., & Trygstad, S. C. (2022). *Strengthening the nordic working life model: Comparative perspectives on institutional resilience and renewal* (Fafo-report No. 2022:13). Oslo: Fafo. Retrieved from https://www.fafo.no/images/pub/2022/10374.pdf

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306–1308. doi: 10.1126/science.adj0998

Eurostat. (2026a). *Artificial intelligence by NACE Rev. 2 activity (isoc_eb_ain2).* Eurostat data browser. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/isoc_eb_ain2/default/table (Data extracted April 2026)

Eurostat. (2026b). *Individuals - use of generative ai tools.* Eurostat data browser. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/isoc_ai_iaiu/default/table?lang=en (Data extracted April 2026)

Fanfani, B. (2023). The employment effects of collective wage bargaining. *Journal of Public Economics*, *227*, 105006. doi: 10.1016/j.jpubeco.2023.105006

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217. doi: 10.1002/smj.3286

Frank, M. R., Sabet, A. J., Simon, L., Bana, S. H., & Yu, R. (2026). *AI-exposed jobs deteriorated before ChatGPT.* arXiv:2601.02554. doi: 10.48550/arXiv.2601.02554

Freund, L. B., & Mann, L. F. (2026). *Job transformation, specialization, and the labor market effects of AI.* (Working paper)

Gambacorta, L., Qiu, H., Shan, S., & Rees, D. M. (2024). *Generative AI and labour productivity: a field experiment on coding* (BIS Working Paper No. 1208). Bank for International Settlements.

Google. (2026). *Google trends: "artificial intelligence", "ai", worldwide, 2015–2026.* https://trends.google.com/explore?q=artificial%2520intelligence%2Cai&date=2015-01-01%202026-04-27&geo=Worldwide. (Accessed: 2026-04-27)

Government of Norway. (2013). *NOU 2013:13 – lønnsdannelsen og utfordringer for norsk økonomi (Holden III).* Official Norwegian Reports, Ministry of Finance. Retrieved from https://www.regjeringen.no/no/dokumenter/nou-2013-13/id747703/

Haapanala, H., Marx, I., & Parolin, Z. (2023). Robots and unions: The moderating effect of organised labour on technological unemployment. *Economic and Industrial Democracy*, *44*(3), 827–852. doi: 10.1177/0143831X221094078

Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., … Ganguli, D. (2025). *Which economic tasks are performed with AI? Evidence from millions of Claude conversations.* arXiv:2503.04761. doi: 10.48550/arXiv.2503.04761

Hosseini, S. M., & Lichtinger, G. (2025). *Generative AI as seniority-biased technological change: Evidence from U.S. résumé and job posting data.* SSRN Electronic Journal. doi: 10.2139/ssrn.5425555

Hui, X., Reshef, O., & Zhou, L. (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*,

35(6), 1977–1989. doi: 10.1287/orsc.2023.18441

Humlum, A., & Vestergaard, E. (2025). *Large language models, small labor market effects* (NBER Working Paper No. 33777). National Bureau of Economic Research. doi: 10.3386/w33777

Iacono, R. (2018). The nordic model of economic development and welfare: Recent developments and future prospects. *Intereconomics*, *53*(4), 185–190.

Kauhanen, A., & Rouvinen, P. (2025). Assessing early labour market effects of generative AI: evidence from population data. *Applied Economics Letters*, 1–4. doi: 10.1080/13504851.2025.2513973

Kauhanen, A., & Rouvinen, P. (2026). *AI has not impacted the youth labor market in Finland* (ETLA Working Paper No. 2026:1). ETLA. doi: 10.2139/ssrn.4904305

Kjellberg, A. (2025). Changes in union density in the Nordic countries. In L. Calmfors, A. Würtz Rasmussen, & R. Uusitalo (Eds.), *Nordic economic policy review 2025: Wage formation and the nordic model* (pp. 123–149). Copenhagen: Nordic Council of Ministers. Retrieved from https://pub.norden.org/nord2025-001/changes-in-union-density-in-the-nordic-countries.html

Koch, M., Manuylov, O., & Smolka, M. (2021). Robots and firms. *The Economic Journal*, *131*(638), 2553–2584. doi: 10.1093/ej/ueab009

Lambert, P. J., & Schindler, Y. (2026, May). *The broken ladder: Ai, remote work, and early-career hiring* (SSRN Working Paper No. 6787638). Social Science Research Network.

Lodefalk, M., Löthman, L., Koch, M., & Engberg, E. (2026). *Same storm, different boats: Generative AI and the age gradient in hiring* (RATIO Working Paper No. 388). RATIO Institute.

Mann, K., & Püttmann, L. (2023). Benign effects of automation: New evidence from patent texts. *Review of Economics and Statistics*, *105*(3), 562–579. doi: 10.1162/rest_a_01083

Massenkoff, M., & McCrory, P. (2026). Labor market impacts of ai: A new measure and early evidence [Note].

Menon Economics. (2025). *The significance of the petroleum industry for the Norwegian economy* (Report commissioned by the Norwegian Ministry of Energy). Menon Economics. Retrieved from https://www.norskpetroleum.no/en/economy/employment/ (Reported via Norwegian Petroleum portal; accessed 2026-04-27)

Mogstad, M., Salvanes, K. G., & Torsvik, G. (2025, September). Income equality in the nordic countries: Myths, facts, and lessons. *Journal of Economic Literature*, *63*(3), 791–839. doi: 10.1257/jel.20251636

Norges Bank. (2021). *Rate decision september 2021.* Monetary Policy Meeting, 22 September 2021. Retrieved from https://www.norges-bank.no/en/topics/Monetary-policy/Monetary-policy-meetings/2021/september-2021/ (Accessed: 2026-04-27)

Norwegian Directorate of Immigration (UDI). (2024). *72,000 refugees from Ukraine have sought protection in Norway.* UDI Statistical Release. Retrieved from https://www.udi.no/en/latest/72000-from-ukraine/ (Accessed: 2026-04-27)

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. doi: 10.1126/science.adh2586

Parolin, Z. (2020a). Automation, occupational earnings trends, and the moderating role of organized labor. *Social Forces*, *99*(3), 921–945. doi: 10.1093/sf/soaa032

Parolin, Z. (2020b). *Organized labor and the employment trajectories of workers in routine jobs: Evidence from U.S. panel data* (Tech. Rep.). Brookings Institution.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). *The impact of AI on developer productivity: Evidence from GitHub Copilot.* arXiv:2302.06590. doi: 10.48550/arXiv.2302.06590

Pizzinelli, C., Panton, A. J., Mendes Tavares, M., Cazzaniga, M., & Li, L. (2023). *Labor market exposure to AI: Cross-country differences and distributional implications* (IMF Working Paper No. 2023/216). International Monetary Fund. doi: 10.5089/9798400254802.001

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, *90*(5), 2555–2591. doi: 10.1093/restud/rdad018

Statistics Norway. (2026). *A-melding (a-ordningen) microdata, 2015–2025.* Statistics Norway. Retrieved from https://www.ssb.no/en/data-til-forskning/utlan-av-data-til-forskere

Teeselink, B. K. (2025). *Generative AI and labor market outcomes: Evidence from the United Kingdom.* SSRN Electronic Journal. doi: 10.2139/ssrn.5516798

Teutloff, O., Einsiedler, J., Kässi, O., Braesemann, F., Mishkin, P., & Del Rio-Chanona, R. M. (2025). Winners and losers of generative AI: Early evidence of shifts in freelancer demand. *Journal of Economic Behavior & Organization*, *235*, 106845. doi: 10.1016/j.jebo.2024.106845

UNHCR. (2024). *Ukraine refugee situation: Operational data portal.* United Nations High Commissioner for Refugees. Retrieved from https://data.unhcr.org/en/situations/ukraine (Gender-disaggregated data on Ukrainian refugees in Europe; accessed 2026-04-27)

U.S. Energy Information Administration. (2024). *Europe Brent spot price FOB (daily).* Petroleum & Other Liquids data series. Retrieved from https://www.eia.gov/dnav/pet/hist/RBRTEd.htm (Accessed: 2026-04-27)

Xu, Y., Zhao, A., & Ding, P. (2024). *Factorial difference-in-differences.* arXiv:2407.11937. doi: 10.48550/arXiv.2407.11937

Yotzov, I., Barrero, J. M., Bloom, N., Bunn, P., Davis, S. J., Foster, K. M., … Wang, B. Z. (2026). *Firm data on AI* (NBER Working Paper No. 34836). National Bureau of Economic Research. doi: 10.3386/w34836

Appendices

A Sample construction

A.1 Monthly panel: main-job selection

This section describes the construction of the monthly panel, which forms the basis of all analyses conducted in this paper (often further aggregated to yearly frequency). This step is necessary, since the raw A-ordningen microdata can contain multiple rows (employment relationships) per worker within a given $(\texttt{year}, \texttt{month})$. We cast this into a single record per $\texttt{worker} \times \texttt{year} \times \texttt{month}$, in line with SSB methodology[9]. This preserves person-level outcomes and covariates, while summarizing across concurrent jobs and consistently choosing a "main job" for categorical attributes.

First, for each job and month, the code constructs an activity indicator as follows:

1. Start from $\texttt{arb_arbmark_status} == 1$ (employee).
2. Exclude non-employment types such as $\texttt{arb_type} \in \{\text{"ikke_ansettelse"}, -1, -2\}$.
3. Require plausible temporal overlap with the month: the contract start $\texttt{arb_start}$ is not after the month end and contract end $\texttt{arb_slutt}$ is not before the month start (missing dates are treated as open-ended).

Then, among *active* jobs for a given person-month, the code selects one *main job* using this lexicographic rule:

$$\text{priority}(\text{row}) = \Big(\mathbf{1}\{\texttt{arb_hovedarbeid} = 1\},\ \texttt{arb_stillingspst},\ \texttt{lonn_fast}\Big),$$

sorted descending. Equivalently: prefer rows flagged as hovedarbeid (main job); tie-break by the largest working time, then by the largest fixed wage. Covariates for the person-month (sex, age, municipality, STYRK-08, NACE, sector, employer IDs, etc.) are taken from this selected main job.

A.2 Sample size and balance across AI exposure quartiles

This section reports the filter pipeline applied to the raw monthly panel before estimation, together with the composition of the universe and the post-filter sample across quartiles of the Massenkoff and McCrory (2026) AI exposure index.

Table 4 traces the size of the worker–month panel through the two main filter steps applied uniformly across specifications: restricting to private-sector firms and to full-time main jobs. The resulting 112.6 million worker–months are then further restricted, depending on

[9] Om datagrunnlaget for registerbasert sysselsettingsstatistikk and Ny metode gir bedre informasjon om arbeidstid

the specification, to (i) workers aged 18–66 with a non-missing 4-digit STYRK occupation (so AI exposure can be assigned), and, for descriptive balance below, (ii) the pre-treatment window January 2015 – October 2022. These additional restrictions reduce the sample to approximately 83.8 million worker–months on roughly 1.4 million unique workers.

Table 4: Sample-selection waterfall (raw monthly panel).

Step	Worker–months
Baseline (raw monthly panel, Jan 2015 – Mar 2025)	302,302,868
after restricting to private-sector firms	175,169,471
after restricting to full-time main jobs	112,593,429

Universe of workers, by AI quartile. Table 5 reports the composition of the universe sample (no analysis filters) across quartiles of the Massenkoff and McCrory (2026) index, with quartile assignment based on each worker's *contemporaneous* 4-digit STYRK occupation. The quartiles differ along observable dimensions: Q3 is older (mean age 44.5) and higher-paid (NOK 69,853 monthly cash wage) on average; Q1 has the highest immigrant share (29%) and the highest male share (64%); Q4 has the youngest mean age aside from Q2 (40.8) and the lowest male share (46%). The sum of unique workers across quartiles (∼5.36 million) exceeds the deduplicated total (3.46 million) because workers can appear in more than one quartile across the years they are observed.

Table 5: Composition of the universe sample by AI exposure quartile (Jan 2015 – Mar 2025).

	Q1 (Low)	Q2	Q3	Q4 (High)	All
Unique 4-digit occupations	100	99	99	99	397
Unique workers	1,220,880	1,413,220	1,259,127	1,466,392	3,464,646
Mean age	40.9	38.0	44.5	40.8	41.2
Share male	0.64	0.49	0.55	0.46	0.53
Share immigrant	0.29	0.21	0.12	0.13	0.18
Mean monthly wage (NOK)	45,489	42,096	69,853	56,299	54,519
Mean firm size	1,479	2,226	3,337	1,195	2,071

Workers passing the analysis filters. Table 6 reports the same statistics for the pre-treatment monthly observations (Jan 2015 – Oct 2022) of workers who pass the four uniform filters described in Section 3: private firms, full-time main job, ages 18–66, and non-missing 4-digit STYRK occupation. Mean wages are computed as $\exp(\overline{\log w})$ on monthly cash wages. The qualitative pattern is preserved: Q3 remains older and higher-paid; Q1 has the largest immigrant share; Q4 contains the largest number of pre-treatment worker–months (29.0 million, against 19.5 million in Q1) reflecting the right-skew of the exposure distribution—that is, more workers are concentrated in the most-exposed occupations than in the least-exposed ones.

The drop in mean firm size between the universe (Table 5) and the post-filter sample (Table 6) follows mechanically from removing public-sector employers.

Table 6: Composition of the post-filter pre-treatment sample by AI exposure quartile (Jan 2015 – Oct 2022).

	Q1 (Low)	Q2	Q3	Q4 (High)	All
Unique 4-digit occupations	100	99	99	99	397
Unique workers	648,506	666,866	454,344	801,583	∼1.4 M
Worker–months	19.5 M	18.2 M	17.1 M	29.0 M	83.8 M
Mean age	39.2	38.0	44.6	41.5	40.8
Share immigrant	0.32	0.24	0.14	0.12	0.20
Mean monthly wage (NOK)	44,558	44,879	73,633	58,217	53,879
Mean firm size	336	395	555	368	408

B AI exposure indices

This appendix provides additional detail on the four occupational AI exposure indices used in the analysis: the AI Occupational Exposure index of Felten et al. (2021), its labor-complementarity adjustment by Pizzinelli et al. (2023), the LLM-task overlap measure of Eloundou et al. (2024), and the usage-weighted hybrid of Massenkoff and McCrory (2026), which builds on the actual-usage data of Handa et al. (2025). The brief description of each index is given in Section 3; here we report summary statistics, pairwise correlations, the distribution of the preferred index, and the most exposed Norwegian occupations.

Summary statistics at the occupation level

Table 7 reports summary statistics for each index after collapsing to the unique 4-digit STYRK occupations to which we successfully merged a US-SOC code (between 396 and 399 occupations, depending on the source). The four indices live on different scales — Felten et al. (2021) delivers a standardized score with mean roughly zero and standard deviation roughly one; Pizzinelli et al. (2023) delivers a positive, bounded score on a wider range; Eloundou et al. (2024) reports a probability-like share of LLM-affected tasks, ranging from 0 to almost 1; Massenkoff and McCrory (2026) reports a usage- and capability-weighted share that is right-skewed with a maximum of 0.49. To make estimates comparable across indices, we standardize each one to mean zero and unit standard deviation in regressions.

Table 7: Summary statistics of AI exposure indices at the 4-digit STYRK occupation level.

Index	Mean	SD	Min	P25	P50	P75	Max	N
Felten et al. (2021)	0.028	0.950	−2.123	−0.812	0.039	0.956	1.488	399
Pizzinelli et al. (2023)	4.451	0.635	0.000	4.124	4.465	4.777	5.862	399
Eloundou et al. (2024)	0.472	0.284	0.000	0.203	0.459	0.749	0.977	396
Massenkoff and McCrory (2026)	0.082	0.106	0.000	0.004	0.032	0.129	0.493	397

Pairwise correlations

Table 8 reports raw Pearson correlations (lower triangle) and Spearman rank correlations (upper triangle) between the four indices at the occupation level. The three task-content-based indices — Felten et al. (2021), Eloundou et al. (2024) and Massenkoff and McCrory

(2026) — correlate strongly with each other (Pearson 0.68–0.90, Spearman 0.79–0.89). The index of Pizzinelli et al. (2023) correlates more weakly with the others (Pearson 0.36–0.46, Spearman 0.44–0.52), reflecting its labor-complementarity adjustment which downweights occupations whose AI-exposed tasks involve high-judgment work.

Table 8: Pairwise correlations between AI exposure indices at the occupation level. Pearson below the diagonal; Spearman above.

	Felten	Pizzinelli	Eloundou	Massenkoff
Felten et al. (2021)	1.00	0.44	0.89	0.79
Pizzinelli et al. (2023)	0.36	1.00	0.47	0.52
Eloundou et al. (2024)	0.90	0.37	1.00	0.85
Massenkoff and McCrory (2026)	0.68	0.46	0.76	1.00

Distribution and most-exposed occupations

Figure 6 plots the distributions of the Massenkoff and McCrory (2026) (left panel) and Eloundou et al. (2024) (right panel) indices across 4-digit STYRK occupations. The Massenkoff distribution is heavily right-skewed: half of the occupations have an exposure score below 0.032, while the most exposed occupation reaches 0.493. By contrast, the Eloundou distribution is more evenly spread across its $[0, 1]$ range, with substantially higher mean (0.47) and median (0.46) exposure scores. Table 9 lists the three most-exposed occupations and their headcount in October 2022 (the last pre-treatment month).

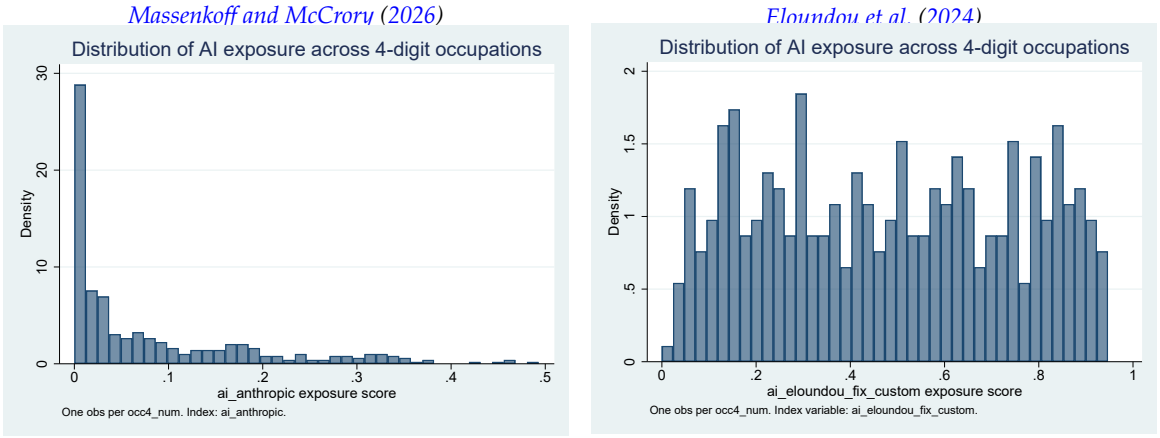


Figure 6: Distribution of the Massenkoff and McCrory (2026) and Eloundou et al. (2024) AI exposure indices across 4-digit STYRK occupations.

Table 9: Three most exposed STYRK occupations under the Massenkoff and McCrory (2026) index (occupations with ≥ 50 workers in October 2022).

Rank	STYRK 4-digit	Massenkoff exposure	Workers (Oct 2022)
1	3324 – Trade and shipping brokers	0.493	583
2	2521 – Database designers, administrators	0.462	361
3	4132 – Data entry clerks	0.456	355

Crosswalk from US SOC to Norwegian STYRK

Each index is published on US Standard Occupational Classification (SOC) codes. We map them to the Norwegian Standard Classification of Occupations (STYRK-08, ISCO-08-based) at the 4-digit level using the SOC–ISCO crosswalks published by O*NET and SSB. Where the mapping is many-to-one (multiple SOC codes feeding into a single STYRK occupation), we use the simple average of the source-SOC scores; where it is one-to-many, we attribute the SOC score uniformly to all matched STYRK codes. Six STYRK occupations could not be merged to any SOC code under either index and are dropped from the AI-exposure analyses; their exclusion absorbs less than 0.5% of worker–months in the universe sample.

C Treatment of concurrent shocks

A number of concurrent shocks possibly affected Norwegian labor markets around the same time as the introduction of GenAI. Most notably, the Covid pandemic disrupted Norwegian labor markets through 2020–2021, with employment losses concentrated among low-paid service workers and parents with young children (Alstadsæter et al., 2020). This asymmetry overlaps with our AI gradient since AI-exposed occupations are typically more teleworkable (Dingel & Neiman, 2020). However, Alstadsæter, Bratsberg, Markussen, Raaum, and Røed (2025) document that by mid-2021 both the level and the socioeconomic composition of Norwegian employment had returned to pre-pandemic patterns for incumbents, with persistent damage confined to those out of work in early 2020. We accordingly extend the pre-period of analysis back to 2015.

A related and more persistent concern is the post-pandemic diffusion of remote work itself. Lambert and Schindler (2026) show that occupational exposure to working from home is strongly correlated with GenAI exposure and argue that the decline in early-career hiring often attributed to GenAI may instead reflect this shift. The timing of our results bears on this: remote work expanded from around 2020, roughly two and a half years before the release of ChatGPT, yet our 22–25 within-firm and shift-share event studies remain essentially flat through 2023 and decline significantly only in 2024, which is hard to reconcile with remote work as the proximate driver. We do not, however, separately identify the two exposures, which are collinear at the occupation level, and we plan to disentangle them directly, for instance by conditioning on a remote-work exposure measure such as Dingel and Neiman (2020), in a future iteration of this paper.

Second, on 22 September 2021, Norges Bank delivered the first post-pandemic policy rate hike among major economies, raising the rate from 0% to 0.25% (Norges Bank, 2021). Lodefalk et al. (2026) flag the analogous Riksbank tightening as a potential confound for AI-period employment trends in Sweden. In Annex E, we decompose the post-period in our firm-composition pooled DiD into a rate-hike indicator (active from September 2021) and a ChatGPT indicator (active from November 2022), so that the rate-hike effect is identified separately from the AI shock. Compared to our main specification, the results do not

materially change.

Third, between 2022 and 2023 Norway received about 72,000 refugees from Ukraine (Norwegian Directorate of Immigration (UDI), 2024), a population that across Europe is roughly 90% women and children due to Ukrainian martial-law restrictions on adult-male emigration (UNHCR, 2024). Our worker-level specification conditions on pre-2022 observed occupations, so Ukrainian refugees will be automatically excluded from the sample, and the outcomes considered there are unlikely to be affected by spillovers due to the compositional effects of their arrival. We further plan to address potential labor-supply effects via robustness checks in the firm-composition and synthetic DiD designs by excluding immigrants altogether in a future iteration of this paper.

Fourth, the war in Ukraine drove Brent oil prices from about $95/barrel pre-invasion to a peak near $139 in mid-2022 (U.S. Energy Information Administration, 2024), with Norwegian petroleum revenues reaching a record NOK 960 billion that year (Menon Economics, 2025). Direct extraction employment is a small share of the Norwegian workforce, though the broader oil-related cluster reaches roughly 10% of private employment with geographical concentration on the west coast (Menon Economics, 2025). Firm-level fixed effects in the within-firm composition design absorb firm-specific responses to oil-related demand.

D Worker-level analysis

Specification for incumbents

We conduct a worker-level analysis on yearly data covering 2015–2024. The panel is restricted to private-sector *full-time* workers aged 18–66 in firms with at least 50 employees,[10] and balanced over the ten-year window: each worker must be observed in every year.[11] The balancing requirement prevents compositional changes such labor-market entry, retirement, and switches to part-time or non-employment, from confounding results. However, this comes at the cost of dropping the youngest cohorts—workers younger than 27 in 2024 cannot have been observed in the panel since age 18 in 2015—so the worker-level estimates below speak to *incumbent* workers; age-specific dynamics are re-examined in the firm- and occupation-level analyses (main text).

We consider four outcomes intended to span both margins of labor-market response: log monthly wages (the intensive-margin price response), months employed in the year (proxying for switches to part-time or unemployment), and binary indicators for changing firm and changing 4-digit occupation between t and $t+1$ (extensive-margin reallocation).

Treatment intensity is the worker's occupation-level AI exposure interacted with a

[10] This is to avoid estimating a large number of firm-year fixed effects, and will be relaxed in a future iteration.

[11] Except for the employment outcome, where we instead balance the panel by filling missing years (unemployment or part-time work) with zeros, but carry AI exposure forward.

post-treatment indicator, so the pooled DiD specification is

$$y_{i,t} = \alpha_i + \delta_{j(i),t} + \beta \, \mathrm{AI}_i \cdot \mathbf{1}[t \geq 2023] + \mathbf{X}'_{i,t}\gamma + \varepsilon_{i,t}, \tag{6}$$

where i indexes workers, t years, $j(i)$ worker i's firm, α_i a worker fixed effect, $\delta_{j(i),t}$ a firm-by-year fixed effect, AI_i occupation-level AI exposure (held fixed at the worker's mean exposure over the period from January 2015 through October 2022), and $\mathbf{X}_{i,t}$ a quadratic in age.[12] Since we run this specification on yearly level, we consider 2022 as the last pre-period year in the pooled specification, and omit it as reference year in the event studies.

The fixed-effect structure is chosen to absorb potential confounders: α_i removes all time-invariant individual heterogeneity (ability, education, gender, immigrant status, baseline wage level), and $\delta_{j(i),t}$ absorbs all firm-level time shocks—including the Norges Bank rate cycle, the oil-revenue boom, or any firm-specific Covid-related dynamics.[13] Identification therefore comes from variation in AI exposure *within* the same firm-year, across workers in different occupations. Standard errors are clustered at the firm level, the unit at which firm-specific demand shocks, hiring policies, and managerial decisions co-move, and the natural level once firm-by-year FE absorb the firm-level shock structure.

This identification strategy makes the worker-level design the most exposed of the four under the factorial parallel trends framework of Section 4: any non-AI shock that loads differentially across occupations within the same firm-year remains unaddressed by the worker and firm-by-year fixed effects. The (conditional) parallel-trends assumption is assessed via event-study plots: pre-treatment coefficients $\hat{\beta}_k$ for $k < 2022$ should be jointly indistinguishable from zero. We complement the visual test with HonestDiD bounds (Rambachan & Roth, 2023) on the post-treatment estimates under specified violations of linear pre-trends.

Lastly, the design is at risk of spillovers: if AI-exposed workers are displaced into low-AI occupations, the comparison group inherits a fraction of the treatment, attenuating $\hat{\beta}$ toward zero. However, for the wage and mobility outcomes the balanced-panel restriction limits this since our sample comprises continuously-employed workers, and the mobility outcomes themselves provide a direct check on the magnitude of any such spillovers.

Incumbent labor-market outcomes

The previous subsection introduced a yearly worker-level difference-in-differences design on a 10-year balanced panel of private-sector employees aged 18–66 in firms with at least 50 employees. The estimating sample for the wage outcome contains 977,967 worker-year observations across 5,793 firms; the mobility outcomes use a near-identical panel of 977,993 observations. The months-employed outcome is estimated on the zero-filled variant of this panel described in Section D.

Table 10 reports the pooled DiD coefficient $\hat{\beta}$ from equation (6) for each of the four

[12] All other observed covariates are time-invariant and hence fully absorbed by unit fixed effects.

[13] The exception is the months-employed outcome, where we replace $\delta_{j(i),t}$ with a year fixed effect δ_t in both equation (6).

outcomes under the Massenkoff and McCrory (2026) index. Three of the four estimates are statistically distinguishable from zero, however these readings turn out to be misleading once the parallel-trends assumption is examined.

Table 10: Pooled difference-in-differences estimates of equation (6) on the worker-level panel (2015–2024). Treatment intensity is the standardized Massenkoff and McCrory (2026) exposure, fixed at each worker's pre-treatment mean.

Outcome	$\hat{\beta}$	SE
Log monthly wage	0.0057^{***}	0.0006
Months employed in year	0.0029	0.0019
Changed firm (next year)	-0.0027^{***}	0.0007
Changed 4-digit occupation (next year)	-0.0042^{***}	0.0011

Standard errors clustered at the firm level (5,793 clusters). All specifications include worker fixed effects, firm-by-year fixed effects (year fixed effects only for the months-employed outcome, see Section D), and a quadratic in age. $^{***}\,p < 0.01$, $^{**}\,p < 0.05$, $^{*}\,p < 0.10$.

Figure 7 plots the event-study coefficients $\hat{\beta}_k$ for each outcome. The wage panel is visibly inconsistent with parallel trends: the pre-2022 coefficients walk monotonically from -0.008 (2015) to -0.002 (2021)—every one of them significantly different from zero at the 1% level—and the post-2022 estimates of $+0.001$ (2023) and $+0.001$ (2024) sit on this seven-year secular trend rather than breaking from it at ChatGPT's release. The pooled coefficient on log wage in Table 10 therefore captures the average level shift between the seven pre-treatment years (which carry mostly negative coefficients) and the two post-treatment years (which carry small positive coefficients), most of which is built up through pre-existing trend rather than a discrete break in 2022. The mobility panels show analogous, if less monotone, pre-period drifts.

Figure 8 reports HonestDiD identified sets (Rambachan & Roth, 2023) for each post-treatment ATT under specified violations of linear pre-trends, parameterised through the standard-deviation grid $M \in \{0, 0.001, \ldots, 0.01\}$. For all four outcomes the identified set already includes zero at $M = 0$ (no allowed violation), and the bounds widen monotonically with M. We therefore cannot reject a zero effect of AI exposure on any of the four worker-level outcomes once linear-pre-trend extrapolations are admitted. Taken together with the visible pre-period drifts, these bounds rule out a causal interpretation of the pooled DiD estimates in Table 10: the apparent wage premium and mobility reductions reflect long-running secular composition trends rather than a discrete labor-market response to the ChatGPT release.

Below we further report a set of robustness checks. Across the seven variants we consider the coefficients shift sign and significance. This instability confirms that we should not interpret any of the worker-level estimates as evidence of a causal AI effect on incumbent workers.

This null at the worker level is in line with expectations, considering that Norway's labor markets are highly regulated; wages are often subject to collective agreements, and labor protection laws restrict layoffs. It however does not rule out any labor-market response

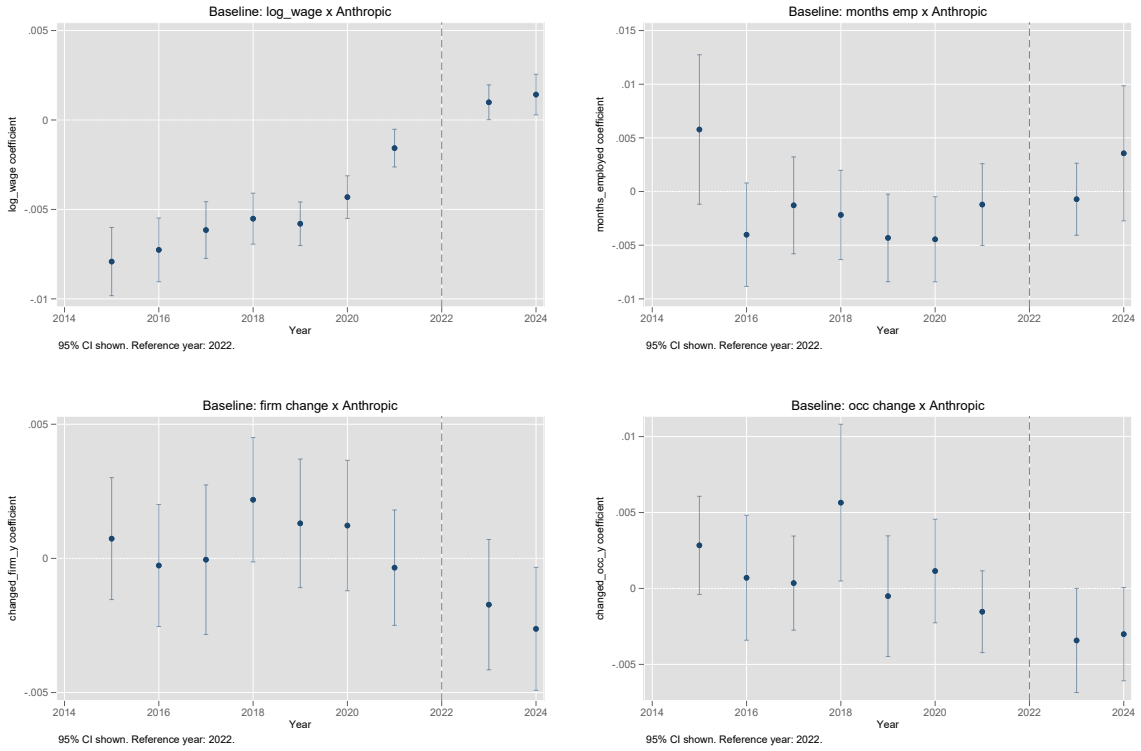


Figure 7: Worker-level event study coefficients $\hat{\beta}_k$. Reference year 2022. Massenkoff and McCrory (2026) exposure index, standardized; treatment intensity fixed at each worker's pre-treatment mean. 95% confidence intervals based on standard errors clustered at the firm level.

to AI exposure in Norway. As Section 5.1 foreshadowed, the within-firm age composition of high-exposure employment did shift after 2022, particularly for younger workers; the design that captures this hiring margin directly is the focus of Section 5.3.

Robustness

Here we report additional robustness checks on the worker-level design, specifically we show (i) the cross-index baseline, and (ii) seven robustness specifications varying the AI-fixing window, treatment parameterization, sample selection, and fixed-effect structure, and (iii) the same baseline event studies estimated on the unbalanced panel.

Cross-index baseline

Table 11 repeats the pooled DiD from Table 10 across the four AI exposure indices. However, relevant event-study plots confirm (strong) pre-trends and Rambachan and Roth (2023) bounds squarely indicate null effects for most specifications.

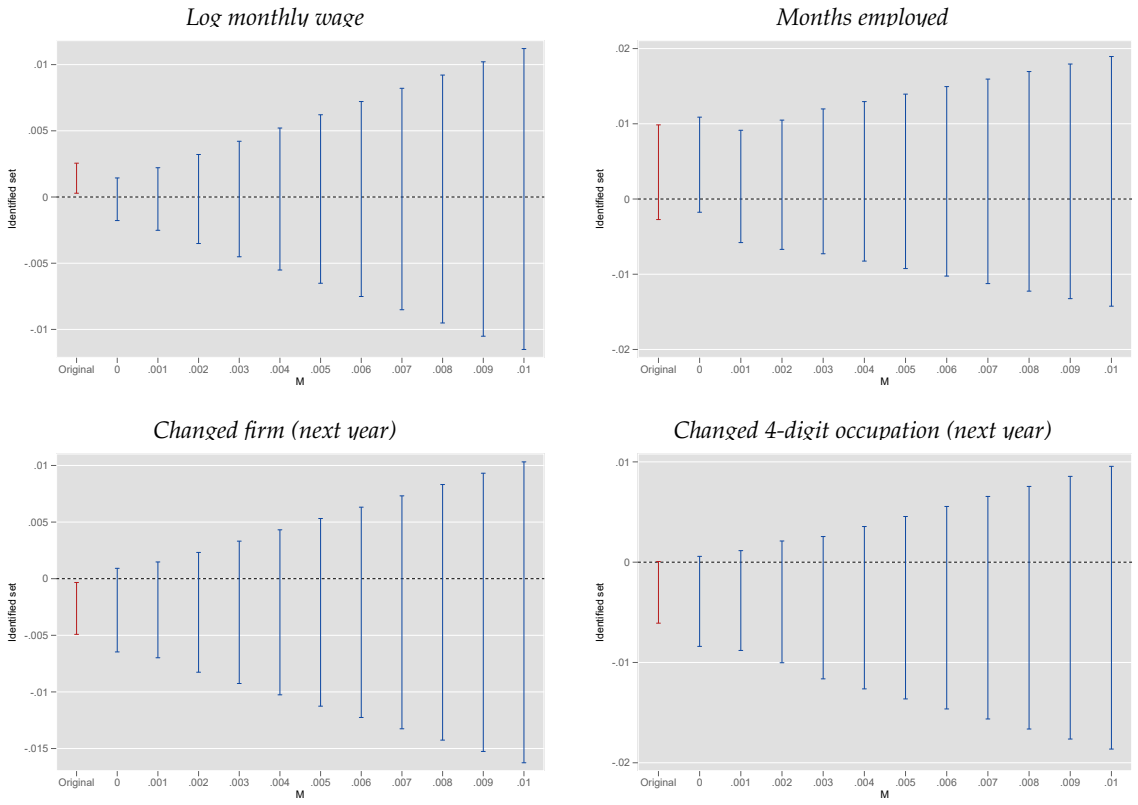


Figure 8: HonestDiD identified sets (Rambachan & Roth, 2023) for the post-treatment ATT in each worker-level outcome, under standard-deviation pre-trend violations $M \in \{0, 0.001, \ldots, 0.01\}$.

Robustness specifications

We run seven robustness variants on the default Massenkoff and McCrory (2026) index. The first four vary the construction of the treatment intensity, the fifth reweights the cross-section, and the last two drop the firm-by-year fixed effect in favour of less restrictive structures.

C1: Modal-2022 fix on a 2020–2024 balanced panel. AI exposure is fixed at the worker's modal occupation in the year 2022 (rather than the pre-treatment-mean used in the baseline), and the panel is rebalanced over 2020–2024 instead of 2015–2024.

C2: Contemporaneous AI. Treatment intensity is the worker's exposure under their *current-year* 4-digit occupation rather than a pre-treatment fix; otherwise identical to the baseline.

C3: Alt-2020 placebo. The treatment year is shifted to 2020 (no AI shock), keeping all other specification elements as in the baseline. A zero placebo coefficient is the test of the design.

C4: Binary high-vs-low. Treatment is discretized as an indicator for above-median AI exposure rather than the standardized continuous score.

Table 11: Pooled DiD coefficient $\hat{\beta}$ from equation (6), by AI index and outcome (worker-level balanced panel, 2015–2024). Standard errors in parentheses, clustered at the firm level.

Outcome	Felten	Pizzinelli	Eloundou	Massenkoff
Log monthly wage	$+0.0165^{***}$	$+0.0027^{***}$	$+0.0146^{***}$	$+0.0057^{***}$
	(0.0009)	(0.0008)	(0.0009)	(0.0006)
Months employed in year	$+0.0011$	$+0.0026$	$+0.0025$	$+0.0029$
	(0.0029)	(0.0028)	(0.0027)	(0.0019)
Changed firm (next year)	-0.0001	-0.0015	-0.0003	-0.0027^{***}
	(0.0010)	(0.0010)	(0.0010)	(0.0007)
Changed 4-digit occupation (next year)	-0.0056^{***}	-0.0037^{**}	-0.0049^{***}	-0.0042^{***}
	(0.0020)	(0.0018)	(0.0018)	(0.0011)

Each cell is one regression. Sample sizes vary marginally with the index used (different missing patterns) but are within ±0.005% of 977,990 worker-years. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

C5: Entropy balance. Treated and control workers are reweighted using entropy balance on pre-2022 means of age, log wage, immigrant share and firm size, before re-estimating the baseline.

C6: Worker + year FE. Drops $\delta_{j(i),t}$ in favour of year fixed effects only; cluster at the 4-digit occupation instead of the firm; firm-size restriction also dropped (a within-firm-by-year identification strategy is no longer the design).

C7: Worker + NACE-by-year FE. Replaces $\delta_{j(i),t}$ with NACE-section-by-year fixed effects; cluster and sample as in C6.

Table 12 reports the pooled DiD coefficient in each variant, for each outcome. Here again results are unstable across specifications.

Worker-level event studies on the unbalanced panel

Figure 9 re-estimates three worker-level event studies on the unbalanced (full-time, private, native, age 18–66) panel rather than on the 10-year balanced panel introduced in the previous subsection. The pre-period drifts appear similar to the baseline specification, with sharp reversals in trends post-treatment. This indicates that compositional effects are at play here, motivating the firm- and occupation-level analyses.

E Within-firm composition: robustness

This appendix expands on the within-firm composition analysis of Section 5.3. We report (i) the pooled DiD on the alternative shorter window of (2019+) and including an additional control for the concurrent rate hiking cycle, as advocated by Lodefalk et al. (2026), (ii) the same age-band decomposition under the Eloundou et al. (2024) task-exposure index, and (iii) the cohort-by-cohort event studies for the 26–30 through 41–49 age bands omitted from the main-text figure.

Table 12: Pooled DiD coefficient $\hat{\beta}$ across robustness specifications, by outcome (Massenkoff and McCrory (2026) index unless noted).

Spec	Log wage	Months emp	Changed firm	Changed occ
Baseline (Table 10)	$+0.0057^{***}$	$+0.0029$	-0.0027^{***}	-0.0042^{***}
	(0.0006)	(0.0019)	(0.0007)	(0.0011)
C1 Modal-2022 (2020–2024)	-0.0016^{**}	$+0.0061^{**}$	-0.0046^{***}	-0.0100^{***}
	(0.0008)	(0.0026)	(0.0013)	(0.0015)
C2 Contemporaneous AI	$+0.0061^{***}$	$+0.0040$	-0.0044^{***}	-0.0073^{***}
	(0.0008)	(0.0025)	(0.0009)	(0.0019)
C3 Alt-2020 placebo	$+0.0008$	$+0.0007$	-0.0015^{***}	-0.0031^{***}
	(0.0005)	(0.0014)	(0.0006)	(0.0010)
C4 Binary above/below median†	$+0.0157^{***}$	-0.0133^{***}	—	—
	(0.0014)	(0.0048)		
C5 Entropy balance	$+0.0015^{***}$	-0.0003	-0.0028^{***}	-0.0041^{***}
	(0.0006)	(0.0018)	(0.0007)	(0.0011)
C6 Spec A (worker + year FE)	-0.0049^{***}	-0.0010	-0.0074^{***}	-0.0047
	(0.0014)	(0.0022)	(0.0025)	(—)
C7 Spec B (worker + NACE×year FE)	$+0.0004$	$+0.0024$	-0.0035^{**}	-0.0055
	(0.0009)	(0.0021)	(0.0016)	(—)

†Coefficient on indicator for "low" (below median) AI; sign therefore inverted relative to continuous specifications. C6 and C7 cluster at the 4-digit occupation (353 / 352 clusters); cells without standard errors arise from a non-symmetric variance matrix on the changed-occupation outcome. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Short window and rate hike robustness

The shorter window restricts the firm panel to cells observed in every month of 2019–2024 (roughly four pre-treatment years), tightening the pre-trend test at the cost of a smaller balanced firm sample. This window is more in line with the shorter studies of Brynjolfsson, Chandar, and Chen (2025) and Lodefalk et al. (2026). In line with the latter, we also additionally control for a rate hike that occurred in 2021, hence adjusting our specification as follows:

$$\log(n_{f,q,t}+1) = \alpha_{f,t} + \gamma_{f,q} + \beta_{\text{NB}}\,\text{Q4}_q \cdot \mathbf{1}[t \geq \text{Sep 2021}] \tag{7}$$

$$+ \beta_{\text{GPT}}\,\text{Q4}_q \cdot \mathbf{1}[t \geq \text{Nov 2022}] + \varepsilon_{f,q,t}, \tag{8}$$

Table 13 repeats the age-band pooled DiD on this window. Compared to the main specification, the results do not materially change.

The qualitative reading from Section 5.3 is preserved on this window: the 22–25 GPT coefficient remains negative, but of smaller magnitude (-0.012) and still insignificant, while the 50+ GPT coefficient remains positive and significant on a dirty pre-trend.

Eloundou index

Table 14 reports the same age-band decomposition on the long (2015+) window under the Eloundou et al. (2024) task-exposure index. The all-ages and 50+ coefficients on $\hat{\beta}_{\text{GPT}}$ are

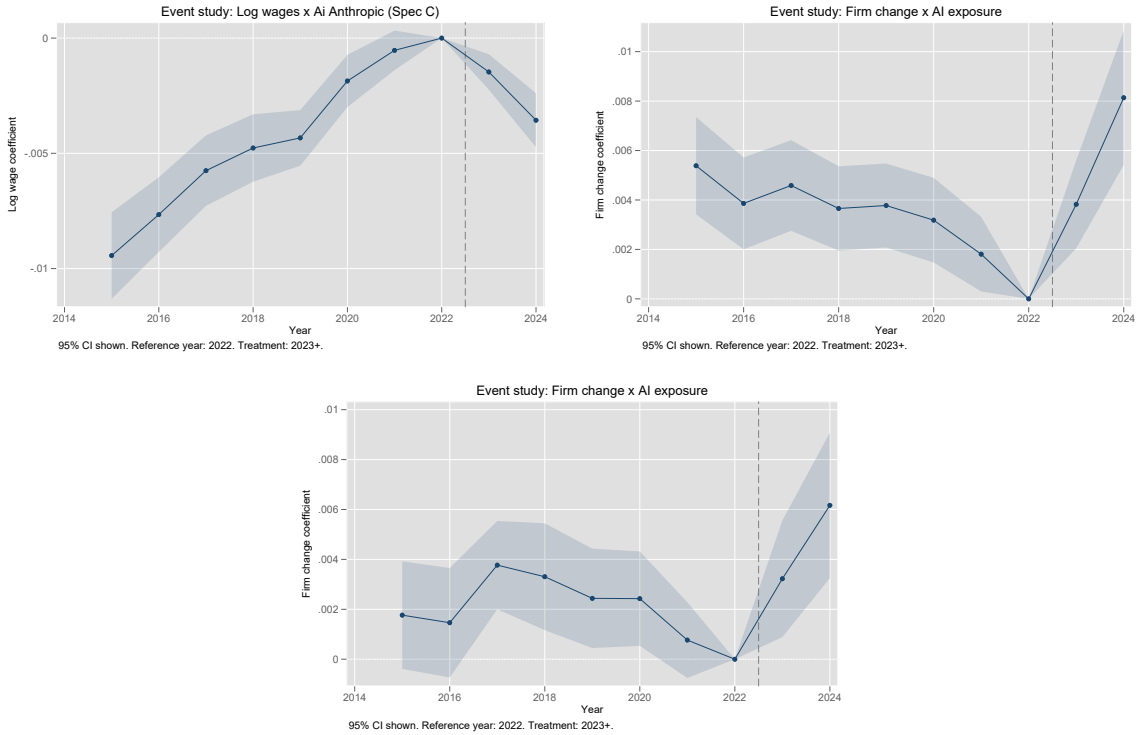


Figure 9: Worker-level event studies on the unbalanced panel (full-time, private-sector natives aged 18–66; balance over 2015–2024 not imposed). Massenkoff and McCrory (2026) exposure index, fixed at each worker's pre-treatment mean. Reference year 2022.

positive and significant, consistent with the Massenkoff results in Table 1. The 22–25 ChatGPT coefficient is again negative but smaller in magnitude (-0.066) and not statistically significant. The pre-trend F is also clean here, but the SE's are wide and sample size is almost as narrow as in the main spec. Cross-index, then, the 22–25 within-firm result is directionally consistent but not conclusively identified under our preferred specification.

Figure 10 plots the half-yearly event-study coefficients of equation (2) under the Eloundou et al. (2024) index for the two age bands of interest, 22–25 and 50+. The 22–25 panel mirrors the Massenkoff result: pre-treatment coefficients are flat (consistent with the clean joint pre-trend test) and post-treatment coefficients drift modestly negative but are not statistically distinguishable from zero. The 50+ panel reproduces the persistent positive trajectory seen under Massenkoff on a similarly dirty pre-trend, flagging the bin as not identified from a parallel-trends standpoint. The cross-index agreement on the qualitative pattern suggests the index choice is not the binding constraint on what we can identify, even if the 22–25 magnitude remains too noisy to support a strong causal reading.

Cohort-by-cohort event studies

Figure 11 reports the half-yearly event-study coefficients of equation (2) for the four age bands omitted from the main-text Figure 3 (26–30, 31–34, 35–40, 41–49). The four bands fill in the age gradient between the marginally clean negative result in 22–25 and the smooth

Table 13: Pooled within-firm DiD coefficients by age band on the short (2019+) window (Massenkoff and McCrory (2026) index).

Age band	$\hat{\beta}_{\text{NB}}$ (SE)	$\hat{\beta}_{\text{GPT}}$ (SE)	Pre-trend $F(6)$	N	Firms
All ages	0.020^{***} (0.003)	0.011^{***} (0.003)	12.27^{***}	3,019,200	10,064
22–25	0.079^{*} (0.045)	-0.012 (0.039)	0.89	30,600	102
26–30	0.079^{***} (0.018)	-0.006 (0.018)	3.25^{***}	145,200	484
31–34	0.040^{**} (0.017)	0.004 (0.015)	3.31^{***}	161,700	539
35–40	0.023^{**} (0.010)	0.013 (0.011)	1.79^{*}	336,000	1,120
41–49	0.010 (0.006)	-0.018^{***} (0.006)	1.88^{*}	713,400	2,378
50+	0.036^{***} (0.004)	0.021^{***} (0.004)	17.89^{***}	1,381,500	4,605

Pre-trend F has 6 numerator degrees of freedom (six pre-2022H1 half-years).

Table 14: Pooled within-firm DiD coefficients by age band on the long (2015+) window, Eloundou et al. (2024) index.

Age	$\hat{\beta}_{\text{GPT}}$	SE	p-value	Pre-trend $F(14)$	N	Firms
All ages	0.0485^{***}	0.0044	0.000	23.08^{***}	3,468,108	7,049
22–25	-0.0662	0.0509	0.195	0.88	27,060	55
26–30	0.0408	0.0267	0.125	1.47	158,916	323
31–34	0.0359	0.0244	0.158	1.63^{**}	177,120	360
35–40	0.0213	0.0165	0.195	1.67^{***}	382,284	777
41–49	-0.0092	0.0109	0.400	2.23^{***}	795,072	1,616
50+	0.0907^{***}	0.0072	0.000	11.76^{***}	1,360,332	2,771

positive trend in 50+: the 26–30, 31–34 and 35–40 panels resemble the 22–25 panel in shape, and the 41–49 panel shows a modest post-2022 dip whose pre-trend test rejects parallel trends at the 5% level.

F Synthetic difference-in-differences: details

This appendix expands on the synthetic difference-in-differences design of Section 5.4. We report (i) the same age-band ATT table under the Eloundou et al. (2024) task-exposure index, (ii) the synthetic-control trajectory plots for the five age bands omitted from the main-text Figure 4, (iii) the unit-weight diagnostics that the Arkhangelsky et al. (2021) estimator returns at each run, (iv) results if the definition of treatment is changed to above-median AI exposure, and (v) results on a publicly available extended occupational level panel that has lower granularity but is updated to include data up to 2026Q1.

Eloundou robustness

Table 15 reports the same age-band SDID ATTs estimated on the Eloundou et al. (2024) index. None of the fourteen real-treatment cells reaches conventional significance under Eloundou; in particular, the 31–34 headcount and wage cells that were significant at the 5% level under Massenkoff and McCrory (2026) (Table 2) become $+0.020$ ($p = 0.29$) and $+0.012$ ($p = 0.21$) respectively, and the 50+ wage cell becomes $+0.008$ ($p = 0.24$). The 22–25 headcount cell is again negative (-0.024) but again not statistically distinguishable

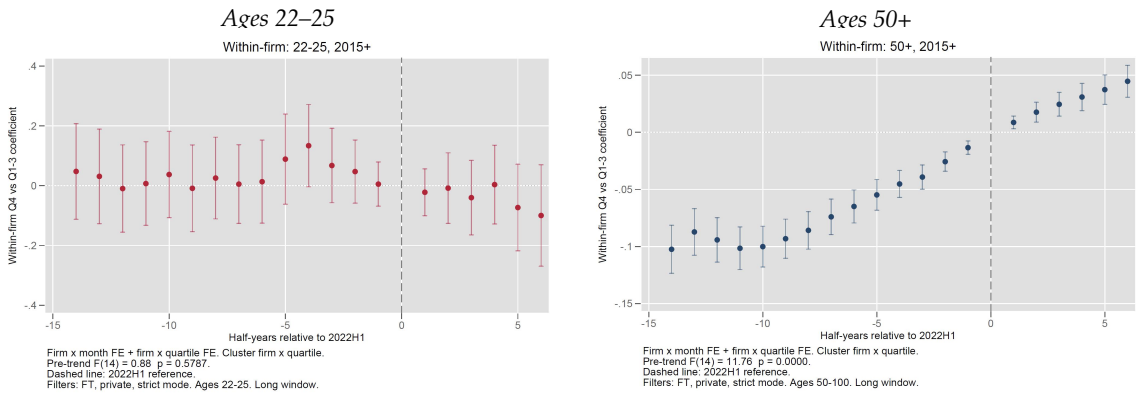


Figure 10: Within-firm composition event-study coefficients $\hat{\beta}_h$ from equation (2) under the Eloundou et al. (2024) index, for the two age bands of interest. Reference period 2022H1. Sample: balanced firm panel (2015–2024). 95% confidence intervals on standard errors clustered at firm $\times$ quartile.

from zero. The cross-index pattern reinforces our reading from the main text: there is no SDID-significant cell that replicates across both indices, so we do not attribute the three Massenkoff-significant cells to a labour-demand response.

Table 15: Synthetic difference-in-differences ATTs at the occupation level, by age band and outcome (Eloundou et al. (2024) index).

	Headcount: $\log(\bar{n}+1)$		Wage: $\overline{\log w}$	
Age band	Real (SE)	2019H2 backdating (SE)	Real (SE)	2019H2 backdating (SE)
All ages	0.000 (0.016)	0.005 (0.023)	0.004 (0.011)	0.016 (0.011)
22–25	−0.024 (0.036)	−0.003 (0.039)	−0.017 (0.015)	−0.004 (0.017)
26–30	0.011 (0.021)	0.055** (0.025)	0.008 (0.009)	0.013 (0.011)
31–34	0.020 (0.019)	0.067** (0.022)	0.012 (0.009)	0.010* (0.006)
35–40	−0.001 (0.015)	0.049** (0.019)	0.004 (0.008)	0.016** (0.005)
41–49	−0.030* (0.015)	0.028* (0.014)	0.006 (0.009)	0.019** (0.005)
50+	−0.010 (0.014)	0.028* (0.015)	0.008 (0.007)	0.010** (0.005)

Column structure as in Table 2. ** $p < 0.05$, * $p < 0.10$.

Trajectories for the remaining age bands

Figures 12 and 13 plot the synthetic-control trajectories for the five age bands omitted from the main-text Figure 4 (all ages, 26–30, 31–34, 35–40, 41–49), separately for headcount and wage. The synthetic-control matches are visibly close to the treated series in every panel through 2022H2; post-treatment divergences are within the bootstrap noise floor, with the 31–34 panels showing the most visible (though still modest) divergence consistent with their nominal significance in Table 2.

Weight diagnostics

The Arkhangelsky et al. (2021) unit weights $\hat{\omega}_o$ that build the synthetic control are returned by the estimator and can be examined directly. Table 16 reports the sum of these weights

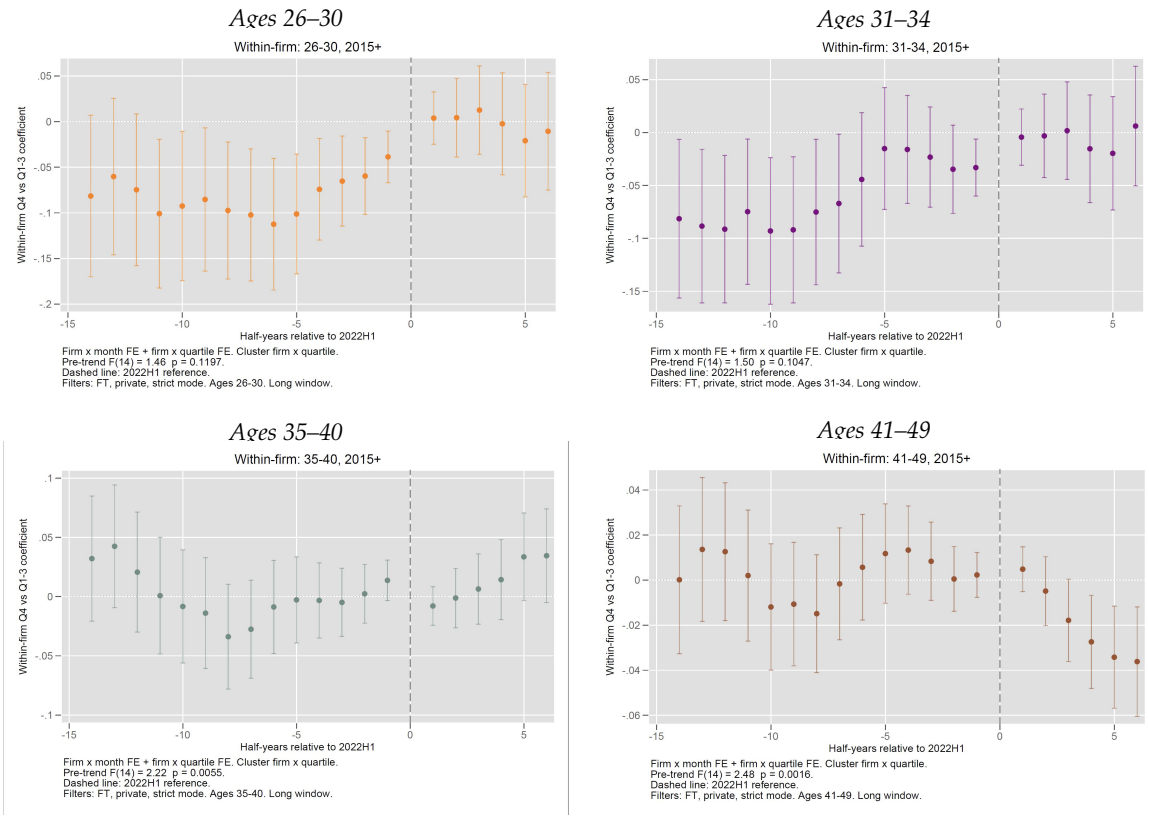


Figure 11: Within-firm composition event-study coefficients $\hat{\beta}_h$ from equation (2), four interior age bands. Reference period 2022H1. Sample: balanced firm panel (2015–2024) with ≥ 1 workers per AI quartile per month; Massenkoff and McCrory (2026) index. 95% confidence intervals on standard errors clustered at firm $\times$ quartile.

by control quartile, separately for each age band and outcome.

Three patterns stand out. First, on the headcount outcome, every age band other than 22–25 puts the most weight on Q3, then Q2, then Q1; the pattern is most pronounced for 31–34, 35–40 and 50+, where Q3 weights reach 0.43–0.45 against Q1 weights of 0.24–0.26. Q3 occupations appear to track the pre-2022 trajectory of Q4 most closely for these older cohorts, consistent with Q3's higher mean wage and older workforce. Second, the youngest age band (22–25) reverses the ranking: Q1 carries the most weight (0.37) and Q3 the least (0.30) on both outcomes, indicating that for younger workers the lowest-exposure occupations provide the closest pre-treatment match. Third, on the mean log wage outcome the cross-quartile weight differences compress substantially—all three quartile sums sit within a 0.05 band of each other in every age bin except 22–25—suggesting that the wage trajectory is less informative than headcount about which controls belong in the synthetic match.

The maximum unit weight assigned to any single control occupation is small in every run (between 0.005 and 0.015 across age bands), so the synthetic match in any given run is constructed from a broad rather than a sparse selection of controls. In this regime the SDID estimate is closer to a quartile-reweighted occupation-level two-way fixed-effects DiD than to a sharp synthetic-control identification, with the cross-quartile reweighting

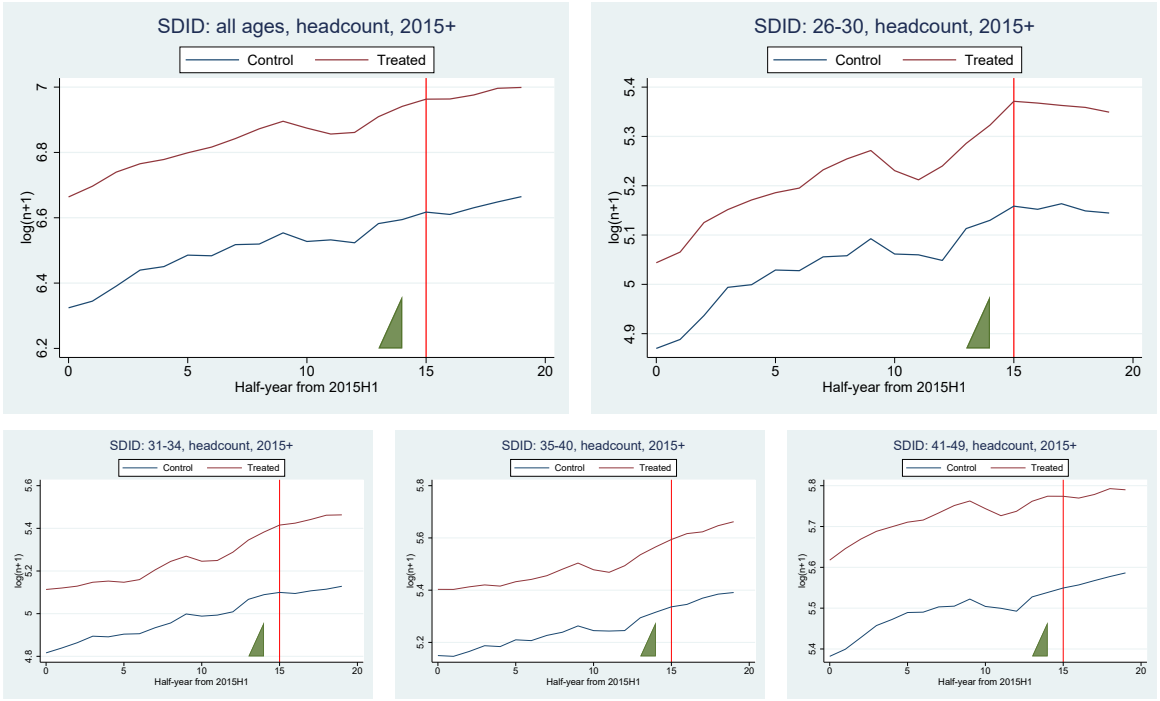


Figure 12: Synthetic-control headcount trajectories, age bands omitted from Figure 4. Vertical reference at 2022H2. Massenkoff and McCrory (2026) index.

and the time-weighting layer doing the bulk of the pre-trend correction. This is consistent with the absence of a visible kink at 2022H2 in the trajectory plots of Figures 4, 12 and 13.

Above-median AI exposure

Table 17 re-runs the SDID design defining the treated set as the above-median half of occupations rather than only the top quartile (Q4), with the below-median half as control. Two readings stand out. First, the 22–25 headcount ATT flips sign from -0.013 in Table 2 to $+0.031$ here, and remains statistically indistinguishable from zero. This is consistent with the descriptive evidence in Figure 1: the post-2022 employment decline among the youngest cohort workers is concentrated in the highest-exposure quartile, while Q1–Q3 remain broadly stable, so pooling Q3 with Q4 in the treated group dilutes the Q4 signal. Second, the 31–34 and 35–40 cells turn positive and highly significant on headcount ($+0.058^{***}$ and $+0.062^{***}$). These positives align with the descriptive pattern in Figure 1, where in these age groups employment in high-exposure occupations grew faster than in low-exposure occupations in particular in Q3. However, these appear to be at least partially driven by pre-existing trends, as the 2019H2 backdating placebos indicate: they are of comparable magnitude to the real-treatment ATTs in most bands, and individually significant for 31–34 headcount ($+0.046^{**}$), 41–49 headcount ($+0.036^{**}$), 50+ headcount ($+0.038^{**}$).



Figure 13: Synthetic-control mean-log-wage trajectories, age bands omitted from Figure 4. Vertical reference at 2022H2. Massenkoff and McCrory (2026) index.

Extended quarterly SSB panel

To test whether the main SDID conclusions are robust to alternative sample definitions and a longer post-period, we re-estimate the design on a publicly available quarterly panel of occupation-level employment and earnings from Statistics Norway.[14]

Two features of this panel are worth highlighting. First, it is available on quarterly basis and the post-period runs through 2026Q1. Second, the SSB tables aggregate over the universe of employees by occupation × age × quarter rather than restricting to private-sector full-time workers, at the cost of coarser age granularity. The most disaggregated band available is "less than 40 years", which pools the 22–25 cohort with the 26–30, 31–34 and 35–40 cohorts, so the SSB panel cannot speak directly to the young-worker displacement that is the focus of the main analysis.

Tables 18 and 19 report the SDID ATTs and the 2019H2 backdating placebos under the Massenkoff and McCrory (2026) and Eloundou et al. (2024) indices respectively, where for this analysis we run the SDID on quarterly instead of half-yearly basis. The pattern is consistent across both indices: real-treatment ATTs are small and positive on headcount in the all-ages and below-40 bands, while the 2019H2 backdating placebos are systematically larger in magnitude and statistically significant in nearly every cell.

This confirms the pre-trend-driven reading of the main results: high-AI occupations

[14]Statistics Norway, Table 11658: "Employees, jobs, and earnings, by sex, age, occupation and quarter", available at https://www.ssb.no/en/statbank/table/11658. The panel covers employees across all occupations and contains 41 quarterly periods from 2016Q1 to 2026Q1. Coverage is the universe of employees by occupation × age × quarter; no full-time or private-sector filter is applied. Age bands are coarser than in our register-based analysis: "all ages", "less than 40 years", "40–54 years", and "55 years or older".

Table 16: Sum of Arkhangelsky et al. (2021) unit weights by control AI quartile, by age band and outcome (Massenkoff and McCrory (2026) index, real-treatment runs).

	Headcount: $\log(\bar{n}+1)$			Wage: $\overline{\log w}$		
Age band	Q1	Q2	Q3	Q1	Q2	Q3
All ages	0.30	0.32	0.39	0.32	0.33	0.35
22–25	0.37	0.33	0.30	0.38	0.35	0.27
26–30	0.30	0.33	0.37	0.31	0.34	0.35
31–34	0.24	0.31	0.45	0.30	0.33	0.37
35–40	0.26	0.29	0.44	0.32	0.33	0.35
41–49	0.29	0.32	0.40	0.31	0.32	0.37
50+	0.25	0.32	0.43	0.30	0.32	0.37

Each cell is the sum of Arkhangelsky et al. (2021) unit weights $\hat{\omega}_o$ across control occupations in that quartile; rows sum to one within each outcome panel. Weights are returned by the estimator and reflect the pre-2022H2 trajectory match against the treated set.

Table 17: Synthetic difference-in-differences ATTs at the occupation level under the above-median treatment definition, by age band and outcome. Massenkoff and McCrory (2026) index.

	Headcount: $\log(\bar{n}+1)$				Wage: $\overline{\log w}$			
Age band	Real (SE)		2019H2 backdating (SE)		Real (SE)		2019H2 backdating (SE)	
All ages	0.012	(0.014)	0.043^{*}	(0.024)	0.005	(0.012)	0.012	(0.011)
22–25	0.031	(0.030)	-0.043	(0.035)	-0.015	(0.016)	0.028	(0.020)
26–30	0.027	(0.018)	0.025	(0.024)	0.012^{*}	(0.007)	0.017^{**}	(0.008)
31–34	0.058^{***}	(0.019)	0.046^{**}	(0.021)	0.014^{**}	(0.007)	0.011^{*}	(0.007)
35–40	0.062^{***}	(0.015)	0.031	(0.020)	-0.001	(0.005)	0.013^{***}	(0.005)
41–49	-0.015	(0.015)	0.036^{**}	(0.017)	0.008	(0.006)	0.012^{*}	(0.007)
50+	0.025^{*}	(0.013)	0.038^{**}	(0.016)	0.009	(0.005)	0.007	(0.007)

Column structure as in Table 2. Treated occupation counts range from 81 (22–25) to 174 (all ages) against 103–173 controls. $^{***}\,p<0.01$, $^{**}\,p<0.05$, $^{*}\,p<0.10$.

have been growing faster than low-AI occupations throughout the pre-period, and continued to do so after 2022 at roughly the same rate. Together with the main-spec results, the SSB panel reinforces the conclusion that the apparent post-2022 employment dynamics in high-AI occupations are largely a continuation of pre-existing trends, and that the magnitude of any clean break attributable to ChatGPT is small, even when extending the analysis to early 2026. However, this analysis cannot speak to the "Canaries cohort" of 22-25 year olds, as here they are pooled together with other age groups that might have exhibited very different employment dynamics.

G Shift-share at the firm level: omitted age bands

This appendix reports the shift-share event-study charts for the five age bands omitted from the main-text Figure 5 (all ages, 26–30, 31–34, 35–40, 41–49), separately under the two AI indices.

Table 18: Synthetic difference-in-differences ATTs at the occupation level on the SSB Table 11658 quarterly panel (2016Q1–2026Q1, treatment 2022Q4), by age band and outcome. Massenkoff and McCrory (2026) index.

	Headcount: $\log(\bar{n}+1)$				Wage: $\overline{\log w}$			
Age band	Real (SE)		2019H2 backdating (SE)		Real (SE)		2019H2 backdating (SE)	
All ages	0.026^{**}	(0.012)	0.045^{***}	(0.015)	0.003^{*}	(0.002)	0.003	(0.002)
Below 40	0.031^{**}	(0.015)	0.044^{**}	(0.017)	0.005^{**}	(0.002)	0.005^{**}	(0.002)
40–54	0.014	(0.011)	0.013	(0.013)	0.007^{***}	(0.002)	0.005^{**}	(0.002)
Above 55	0.019^{*}	(0.011)	0.039^{***}	(0.014)	0.004^{*}	(0.002)	0.005^{**}	(0.002)

Standard errors via placebo bootstrap (500 replications). $^{***}\,p<0.01$, $^{**}\,p<0.05$, $^{*}\,p<0.10$.

Table 19: Synthetic difference-in-differences ATTs at the occupation level on the SSB Table 11658 quarterly panel, by age band and outcome. Eloundou et al. (2024) index. Format as in Table 18.

	Headcount: $\log(\bar{n}+1)$				Wage: $\overline{\log w}$			
Age band	Real (SE)		2019H2 backdating (SE)		Real (SE)		2019H2 backdating (SE)	
All ages	0.024^{*}	(0.014)	0.058^{***}	(0.014)	0.003^{*}	(0.002)	0.005^{***}	(0.002)
Below 40	0.024	(0.016)	0.067^{***}	(0.015)	0.005^{*}	(0.002)	0.005^{**}	(0.002)
40–54	0.013	(0.013)	0.036^{***}	(0.011)	0.008^{***}	(0.002)	0.007^{***}	(0.002)
Above 55	0.009	(0.012)	0.052^{***}	(0.011)	0.004	(0.002)	0.005^{***}	(0.002)

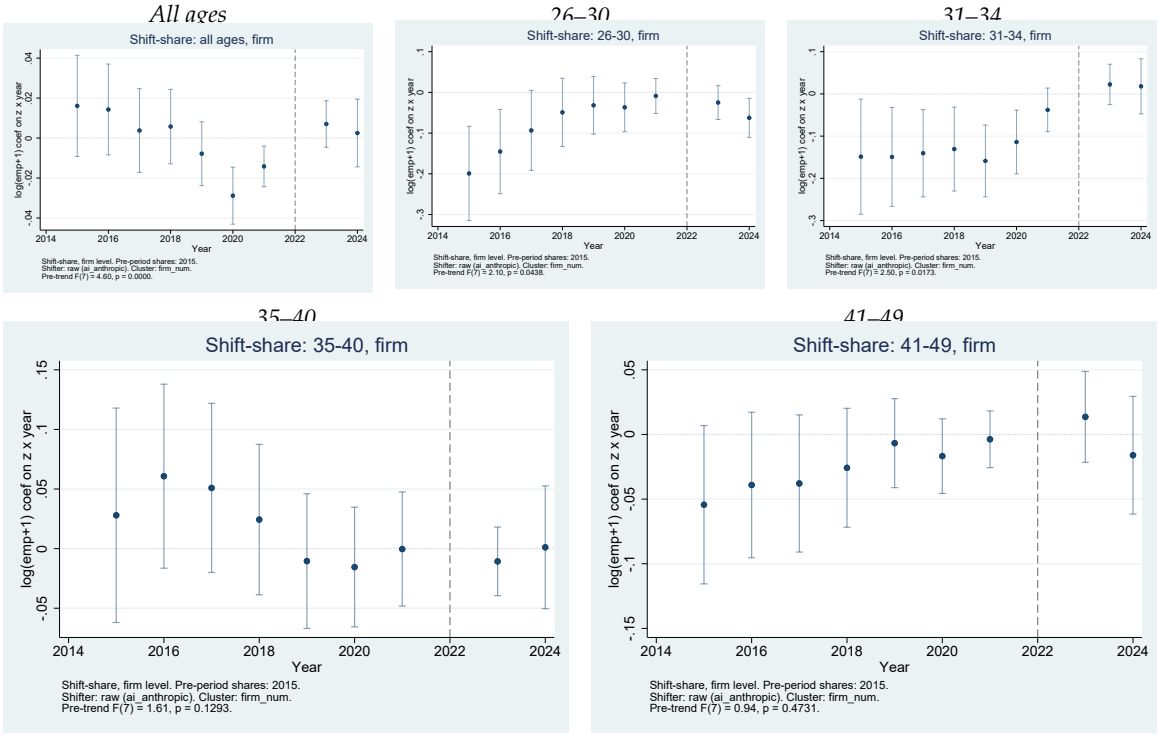


Figure 14: Shift-share event-study coefficients $\hat{\beta}_{a,k}$ from equation (5) for the five age bands omitted from the main-text Figure 5, Massenkoff and McCrory (2026) index. Reference year 2022. 95% confidence intervals based on standard errors clustered at the firm level.

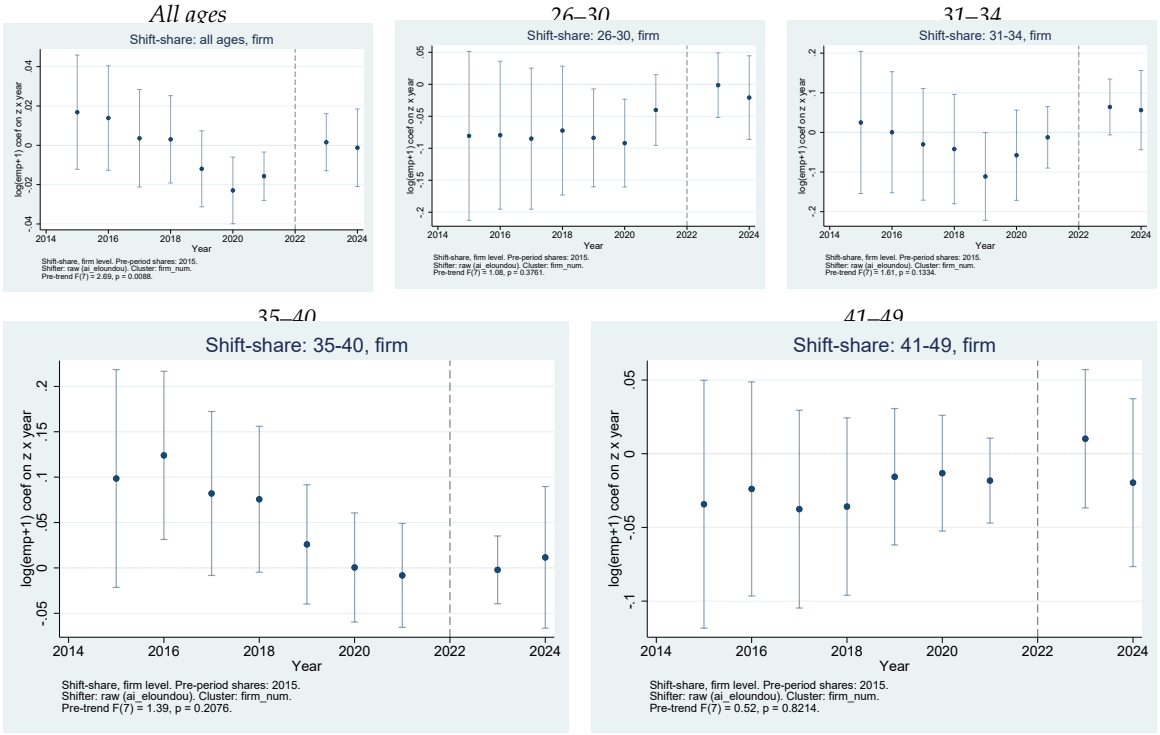


Figure 15: Shift-share event-study coefficients $\hat{\beta}_{a,k}$ from equation (5) for the five age bands omitted from the main-text Figure 5, Eloundou et al. (2024) index. Reference year 2022. 95% confidence intervals based on standard errors clustered at the firm level.