International Labour Organization

ILO brief

Research Brief

May 2026

The impact of GenAI on jobs, productivity and work organization: a review of the empirical evidence[1]

This article synthesises empirical evidence on how Generative AI is affecting productivity, employment, and work organisation. Drawing on experiments, firm-level data, platform studies and new representative worker and firm surveys from Australia, Denmark, Germany, Korea, Kuwait, the United Kingdom and the United States, it finds that productivity gains are real albeit often unverified and uneven. Large-scale job displacement remains limited, and worker-reported time savings of a few per cent of working hours have not yet translated into higher measured output, earnings or employment. The main risks lie in growing inequalities, the erosion of employment opportunities for younger workers, and the reshaping of how work is organised, with implications for coordination, worker autonomy, and job quality.

Introduction

Generative artificial intelligence (GenAI) is rapidly diffusing across workplaces, reshaping how tasks are performed and how work is organised. While ex-ante AI exposure indicators provide insights into which jobs may be affected, emerging empirical evidence allows policymakers to assess how GenAI has started to influence labour markets in practice.

Drawing on evidence from experiments, firm-level data, digital labour platforms, and organisational studies, this Research Brief synthesises key findings from Del Rio Chanona et al. (2025) regarding productivity and employment effects of generative AI.

There are three central messages. First, GenAI can generate substantial productivity gains, but these gains are distributed unevenly across tasks, workers, and organisational contexts. Second, GenAI might amplify inequalities in performance and access to job opportunities and reshape work organisation in ways that may undermine productivity and job quality. Third, the productivity gains documented at the task level have not yet shown up as higher measured output, earnings or working hours, either in representative worker surveys (Denmark, Korea, the United States) or in firm-level surveys covering four advanced economies (Australia, Germany, the United Kingdom, the United States). Workers appear to capture the time savings primarily as on-the-job leisure and as reallocation toward tasks the technology does not affect, rather than as expanded output, work pressure reductions or pay increases. This gives rise to an "aggregation paradox" (Chan and Shedania, 2026). Finally, GenAI is beginning to reshape work organisation itself. Evidence suggests that AI affects not only individual tasks but also the organisation of work as bundles of interdependent activities. Rather than eliminating entire occupations, AI mainly reorganises tasks within jobs, with workers reallocating effort towards tasks less easily displaced by automation. In some settings, AI assistance allows less-

[1] By Rossana Merola (ILO), Ekkehard Ernst (ILO), Daniel Samaan (ILO), Maria del Rio-Chanona (University College London), Ole Teutloff (University of Oxford).

experienced workers to perform tasks previously reserved for specialists, flattening skill hierarchies and changing team composition. At the same time, greater reliance on AI-generated suggestions may reduce diversity of ideas and lead to more standardised outputs.

▶ Productivity effects

A rapidly expanding empirical literature reviewed in Del Rio Chanona et al. (2025) shows that GenAI can raise productivity, but that the size and distribution of gains depend strongly on task scope, task complexity, and how AI is integrated into workflows (see table 1 and figure 1 for a summary overview of the empirical evidence).

In controlled Randomised Controlled Trials (RCTs)[2] that focused on relatively simple tasks, large time savings are common in writing and coding: mid-level writing tasks are completed substantially faster with AI support (Noy and Zhang, 2023), and similar patterns appear in professional and creative writing (Li et al., 2024) as well as legal drafting exercises (Choi et al., 2024). However, much of this gain in speed may reflect passive pasting, workers submitting AI output with minimal modification, suggesting automation rather than skill enhancement, particularly for lower-skilled workers where AI performance already meets task requirements.

In programming, Copilot-style assistance can substantially speed up narrowly defined tasks (Peng et al., 2023). The evidence also shows important caveats: for creative or expert-intensive writing, AI can reduce originality or harm expert performance through anchoring and reduced diversity (Chen and Chan, 2024; Doshi and Hauser, 2024), and in mathematical problem-solving, assistance can lower performance when users cannot detect errors or hallucinations (Kim and Moon, 2024).

Beyond text and code, classification experiments show that human–AI collaboration is not automatically beneficial: in radiology, providing AI predictions does not necessarily improve diagnostic accuracy (Agarwal et al., 2023), while in image-based judgments, AI can raise accuracy, but effects depend on users 'calibration and self-assessment (Caplin et al., 2024). Evidence from debate preparation similarly suggests that benefits can concentrate among top performers rather than equalising outcomes (Roldán-Monés, 2024).

For complex tasks in controlled settings, the literature remains smaller but provides more nuanced lessons about when productivity gains translate into better decisions. Studies covering consulting, academic writing, general knowledge work and legal work show that GenAI can improve speed and sometimes quality, but effects are more sensitive to ambiguity, context, and the user's ability to judge when AI is unreliable (Dell'Acqua et al., 2023; Usdan et al., 2024; Haslberger et al., 2025; Schwarcz et al., 2025). A recurring mechanism is the "jagged technological frontier": performance improves when tasks are within a model's capability frontier, while tasks outside the frontier can generate misleading outputs and raise error risks, especially under over-reliance (Dell'Acqua et al., 2023; Agarwal et al., 2023).

Field experiments confirm this pattern: productivity impacts depend on both the task type, and the structure of AI integration. In telemarketing, sequential division of labour—AI handling lead generation while humans handle persuasion—raises performance for higher-skilled agents but can disadvantage lower-skilled agents by leaving them with a higher concentration of difficult cases (Jia et al., 2024). In customer support, real-time conversational assistance boosts productivity and quality most for less-experienced workers, narrowing performance gaps (Brynjolfsson, Li and Raymond, 2025). At a broader organisational scale, early evidence from Microsoft 365 Copilot adoption across many firms suggests meaningful time savings on routine individual tasks (notably email and document drafting), but limited signs of deeper organisational transformation in the short run (Dillon et al., 2025). For more complex field settings, results are heterogeneous: among entrepreneurs, average effects may be small while high performers benefit and low performers can be harmed, reflecting differences in how users translate generic AI suggestions into context-specific actions (Otis et al., 2023).

Field and quasi-experimental evidence on software development reveals highly context-dependent

[2] Randomised Controlled Trials (RCT) is a research method used to identify causal effects by comparing outcomes between two (or more) groups that are formed randomly: a treatment group (e.g. workers who are given access to GenAI tools), and a control group (e.g. similar workers who are not given access). Because assignment is random, differences in outcomes—such as productivity, quality of output, or time spent on tasks—can be attributed with high confidence to the intervention itself (here, GenAI use), rather than to pre-existing differences between workers.

productivity effects, with gains varying substantially by developer experience and workplace setting. In a large multi-firm experiment with enterprise developers, AI coding assistants increased completed tasks, with larger gains among less-experienced developers (Cui et al., 2025). However, a controlled experiment with experienced open-source developers found that AI tools increased task completion time by 19 percent (Becker et al., 2025). Exploiting the temporary ChatGPT ban in Italy, one study finds no overall productivity changes for experienced developers but minor slowdowns in debugging tasks, while less-experienced developers increased both output quantity and quality when AI access was *removed* (Kreitmeir and Raschky, 2024). A study using GitHub data finds modest productivity gains averaging 2-4 percent, concentrated among experienced developers, with early-career developers showing no significant benefits (Daniotti et al., 2026).

Beyond productivity, research on Copilot adoption shows that individual developers reallocate time away from project management toward autonomous coding (Hoffmann et al., 2025), but increased code contributions raise team-level coordination costs through more extensive discussions (Song et al., 2023), with new contributions concentrated in maintenance rather than novel features (Yeverechyahu et al., 2024).

Recent representative-survey evidence helps to put these experimental findings into macroeconomic context and to clarify where the time savings are going. In a nationally representative survey of 5,512 Korean workers, Suh and Oh (2026) find that 51.8 per cent of workers use GenAI for their job, saving an average of 3.8 per cent of active working time, equivalent to 1.4 per cent of working hours across the full workforce. The weighted correlation between self-reported time savings and self-reported output changes is low, however: workers save time, but they do not on average produce more. The Real-Time Population Survey in the United States produces strikingly similar magnitudes (Bick, Blandin and Deming, 2025), with users reporting savings of 5.4 per cent of weekly working hours and again 1.4 per cent of total working hours assisted by GenAI economy-wide. Danish administrative records yield self-reported time savings of about 3 per cent of hours but no detectable effect on either earnings or recorded hours within two years of ChatGPT's launch (Humlum and Vestergaard, 2025). In Korea, the time savings are absorbed primarily as a 1.3 percentage-point rise in the share of time spent on on-the-job leisure and as task reallocation, with a single top-saving task accounting for roughly 70 per cent of a user's total time reduction. These patterns are consistent with the workplace-level evidence reviewed above, where individual workers save time on email and routine drafting (Dillon et al., 2025) and developers shift effort from coordination toward core technical work (Hoffmann et al., 2025), without changes in meetings, team output, or measured productivity. Moreover, a key challenges for companies to reallocate these gains in time savings is the poor measurement and lack of key performance indicators especially among small and medium-sized enterprises (ILO, 2026).

Taken together, the evidence supports a clear message: the effect of GenAI on productivity depends on the degree of task complexity, on worker expertise, and on the implementation context. GenAI can raise productivity substantially in well-bounded tasks, but many gains from simple tasks reflect passive automation rather than skill enhancement. In complex work, outcomes hinge on workers' ability to judge AI reliability and avoid over-reliance. Crucially, inequality effects vary by context: AI tends to narrow performance gaps in simple tasks but can widen them in complex, judgment-intensive work, particularly when users lack the expertise to validate AI outputs.

▶ Labour demand and employment effects

While productivity effects are increasingly well documented, evidence on labour demand and employment remains more limited and context specific. Overall, there is little evidence so far of large-scale job losses directly attributable to GenAI. However, early signals point to important adjustments in task demand, hiring patterns, and the structure of employment, particularly in occupations heavily reliant on routine cognitive tasks.

Digital trace data from online labour platforms provide some of the clearest evidence of substitution effects. Studies using job postings and online labour platform data show that, following the diffusion of GenAI tools, demand for skills associated with routine writing and translation declined significantly (Liu et al. 2025; Qiao et al., 2024; Yilmaz et al., 2023), while demand for AI-related skills such as machine learning and natural language processing increased (Teutloff et al., 2024). Emerging evidence also points to declining demand in software and web development tasks, particularly for more routine and entry-level work (Demirci et al., 2025).

Evidence from China provides an additional perspective on how GenAI may affect labour demand in different economic contexts. Chen et al. (2025) show that exposure to GenAI varies widely across occupations, and that firms respond primarily by reorganising tasks rather than reducing headcount, again pointing to task-level rather than occupation-level adjustment. This evidence reinforces the view that GenAI affects labour demand by reshaping task bundles and skill requirements rather than by eliminating entire occupations.

While some studies suggest declining employment for entry-level workers in AI-exposed occupations, the emerging consensus points to limited robust evidence for widespread displacement effects. Teutloff et al. (2024) document reduced demand for less-experienced freelancers in AI-exposed tasks on online platforms, though this represents a highly flexible market segment where task-level substitution may occur more rapidly than in traditional employment. Potentially the strongest evidence for broader effects comes from Klein Teeselink (2025), who finds that UK firms with high AI exposure reduced employment in junior positions. However, the study acknowledges important limitations: employment data derived from LinkedIn likely overrepresents professional workers and the 30-month observation window makes it difficult to rule out other concurrent shocks. In the US, Brynjolfsson et al. (2025) report a 16 percent relative decline in employment for workers aged 22–25 in AI-exposed occupations, but this finding faces methodological critiques: the timing precedes ChatGPT's launch and aligns with monetary policy tightening in interest rate-sensitive sectors, suggesting macroeconomic rather than technological drivers (Frank et al., 2026; Iscenko and Millet, 2026). In contrast, comprehensive administrative data covering entire labour markets find no employment effects: Denmark's detailed registers show effectively zero impact on earnings, hours worked, or net job creation from AI adoption through 2024 (Humlum and Vestergaard, 2025).

Overall, aggregate employment effects remain modest despite rapid AI diffusion. Corporate AI adoption remains limited: US Census data shows only around 12 percent of large businesses reported using AI for production by late 2025 (U.S. Census Bureau, 2025). This gap likely reflects implementation barriers including regulatory uncertainty, reliability concerns, and the need for complementary investments in infrastructure and organizational redesign to realize productivity gains (Brynjolfsson et al., 2021). As with earlier general-purpose technologies, full labour market effects may take years to materialize. These findings are consistent with broader evidence on the aggregation of AI productivity gains: while task-level gains are large and well-documented, they have yet to translate into measurable macroeconomic productivity growth, reflecting slow diffusion, limited complementary investment, and the historical pattern of delayed returns from general-purpose technologies (Chan and Shedania, 2026).

Firm-side evidence corroborates these worker-level patterns and extends them to economies for which representative AI data had been scarce. Yotzov et al. (2026) field the first representative international firm-level survey on AI use, covering nearly 6,000 CFOs, CEOs and senior executives across the United States, the United Kingdom, Germany and Australia. About 70 per cent of firms report active AI use, concentrated in younger and more productive firms, yet more than 80 per cent of these same firms report no impact of AI on either employment or productivity over the past three years. Average use among top executives is only about 1.5 hours per week, with a quarter reporting no use at all. Looking ahead, executives forecast that AI will boost productivity by 1.4 per cent and cut employment by 0.7 per cent over the next three years, while employees surveyed in parallel forecast a 0.5 per cent net employment increase over the same horizon, amounting to a 1.2 percentage-point gap in employment expectations between those who decide on adoption and those who experience it.

Beyond individual tasks: GenAI and work organization

Del Rio Chanona et al. (2025) highlight that GenAI's effects extend beyond individual productivity to the organisation of work itself. Increasingly, AI tools are embedded within teams, workflows, and management processes, shaping coordination, decision-making, and collective performance.

A growing set of experimental and organisational studies demonstrates that GenAI can improve the speed, consistency, and average quality of outputs produced by teams, particularly in knowledge-intensive activities such as brainstorming, coding, marketing, and document drafting (Dell'Acqua et al., 2023; Brynjolfsson, Li and Raymond, 2025). By providing rapid access to information and standardised suggestions, GenAI can reduce cognitive bottlenecks and lower coordination costs within teams. However, these same mechanisms also introduce important trade-offs. Several studies show that reliance on AI-generated suggestions can reduce the diversity of ideas and lead to convergence toward average solutions, potentially constraining creativity and innovation in contexts where heterogeneity of perspectives is essential (Dell'Acqua et al., 2023).

Evidence further suggests that GenAI alters how expertise is deployed within organisations. In some settings, AI assistance allows less-experienced workers to perform tasks previously reserved for specialists, flattening skill hierarchies and changing team composition (Brynjolfsson, Li and Raymond, 2025). In other contexts, AI shifts the focus of expert work away from implementation toward supervision, verification, and judgment, increasing the importance of workers 'ability to evaluate and contextualise AI outputs (Dell'Acqua et al., 2023). These dynamics imply that organisational gains from GenAI depend not only on access to the technology but also on how roles and responsibilities are redesigned.

A growing literature emphasises that GenAI affects not only individual tasks but the organisation of work as bundles of interdependent activities. Occupations are not collections of separable tasks, and task-level automation does not translate directly into job displacement. Foundational work shows that skills and tasks are bundled and interact through complementarities (Edmond and Mongey, 2021).

Recent contributions show that AI primarily operates through within-job reorganisation. Workers reallocate effort towards not displaced tasks within the same job (Hampole et al., 2025). Whether jobs are displaced depends on the cost of unbundling: "strong" bundles tend to persist and be reorganised rather than replaced (Garicano, Li and Wu, 2026), while organisational structure and knowledge hierarchies also shape outcomes, as jobs are positions in hierarchies that reorganise rather than dissolve under AI (Ide and Talamàs, 2025). AI induces job transformation through changes in task composition, with heterogeneous effects across workers (Freund and Mann, 2026). From a longer-term perspective, occupations evolve as expanding bundles of tasks, suggesting jobs might be more resilient to automation as they gradually include new tasks reliant on human input (Jackson, 2025). Taken together, viewing jobs as task bundles that cannot easily be undone suggests that the impact of artificial intelligence on jobs and productivity might be multi-faceted and depend on factors other than the technology alone.

▶ References

Agarwal, N. et al. (2023). *Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology*, NBER Working Paper 31422 (Cambridge, MA, National Bureau of Economic Research).

Becker, J., Rush, N., Barnes, B., & Rein, D. (2025). Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. Model Evaluation & Threat Research (METR).

Bick, A.; Blandin, A.; Deming, D. J. (2025). *The Rapid Adoption of Generative AI*, NBER Working Paper 32966 (Cambridge, MA, National Bureau of Economic Research); forthcoming, Management Science (2026).

Brynjolfsson, E.; Chandar, A.; Chen, R. (2024). *AI Adoption and Hiring Patterns in U.S. Firms*, Technical Report (Stanford, Stanford Digital Economy Lab).

Brynjolfsson, E.; Li, D.; Raymond, L. (2025). "Generative AI at Work", in *The Quarterly Journal of Economics*, vol. 140, no. 2, pp. 889–942.

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. American Economic Journal: Macroeconomics, 13(1), 333-372.

Caplin, A. et al. (2024). *The ABCs of Who Benefits from Working with AI: Ability, Beliefs, and Calibration*, NBER Working Paper 33021 (Cambridge, MA, National Bureau of Economic Research).

Chan, C. Y. C.; Shedania, K. (2026). *The Aggregation Paradox of AI: Why do micro-economic productivity gains from AI disappear at scale*, Geneva: International Labour Office. https://doi.org/10.54394/00034342

Chen, Q. et al. (2025). "Large Language Models at Work in China's Labor Market", in *China Economic Review*, vol. 92.

Chen, Z.; Chan, J. (2024). "Large Language Models in Creative Work: The Role of Collaboration Modality and User Expertise", in *Management Science*, vol. 70, no. 12.

Choi, J. et al. (2024). "Lawyering in the Age of Artificial Intelligence", in *Minnesota Law Review*, vol. 109.

Cui, Z. K. et al. (2025). "The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers", in *SSRN Electronic Journal*.

Daniotti, S. Et al. (2026). Who is using AI to code? Global diffusion and impact of generative AI. *Science*.

Dell'Acqua, F. et al. (2023). *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity*, NBER Working Paper 31196 (Cambridge, MA, National Bureau of Economic Research).

Del Rio Chanona, R. M.; Ernst, E.; Merola, R.; Samaan, D.; Teutloff, O. (2025). *AI and Jobs: A Review of Theory, Estimates, and Evidence*. Available at: https://arxiv.org/abs/2509.15265.

Demirci, O.; Hannane, J.; Zhu, X. 2025. "Who is AI replacing? The impact of generative AI on online freelancing platforms", in *Management Science*.

Dillon, E. W. et al. (2025). *Shifting Work Patterns with Generative AI*, Technical Report (Cambridge, MA, National Bureau of Economic Research).

Doshi, A. R.; Hauser, O. P. (2024). "Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content", in *Science Advances*, vol. 10, no. 28, p. eadn5290.

Edmond, C.; Mongey, S. (2021). *Unbundling Labor*. Working Paper (University of Melbourne; University of Chicago; NBER; Wharton Finance Working Paper).

Frank, M. R. et al. (2026). AI-exposed jobs deteriorated before ChatGPT. arXiv:2601.02554v1.

Garicano, L.; Li, J.; Wu, Y. (2026). *Weak Bundle, Strong Bundle: How AI Redraws Job Boundaries*. Working Paper.

Hampole, M.; Papanikolaou, D.; Schmidt, L. D. W.; Seegmiller, B. (2025). *Artificial Intelligence and the Labor Market*. NBER Working Paper No. 33509; SSRN 5121025.

Haslberger, M.; Gingrich, J.; Bhatia, J. (2025). "No Great Equalizer: Experimental Evidence on Productivity Effects of Generative AI Use in the UK Labor Market", in *SSRN Electronic Journal*.

Hoffmann, M. et al. (2025). *Generative AI and the Nature of Work*, Harvard Business School Strategy Unit Working Paper 25-021; SSRN 500708.

Humlum, A.; Verstergaard, E. (2025). *Large Language Models, Small Labor Market Effects*, NBER Working Paper 33777 (Cambridge, MA, National Bureau of Economic Research).

Ide, E.; Talamàs, E. (2025). "Artificial Intelligence in the Knowledge Economy", in *Journal of Political Economy*, vol. 133, no. 12.

ILO. (2026). *Adoption of Artificial Intelligence and productivity growth in Kuwait's private sector*. Geneva: International Labour Office. https://doi.org/10.54394/00033397

Iscenko, Z. & Curto Millet, F. (2026). Looking for the Ladder: Is AI Impacting Entry-Level Jobs? Economic Innovation Group.

Jackson, M. (2025). *The Division of Rationalized Labor*. Cambridge, MA: Harvard University Press.

Jia, N. et al. (2024). "When and how artificial intelligence augments employee creativity", in *Academy of Management Journal*, vol. 67, no. 1, pp. 5–32.

Klein Teeselink, B. (2025). "Generative AI and labor market outcomes: Evidence from the United Kingdom." *Available at SSRN*.

Kim, D. G.; Moon, A. (2024). "From Helping Hand to Stumbling Block: The ChatGPT Paradox in Competency Experiment", in *Applied Economics Letters*, pp. 1–5.

Kreitmeir, D.; Raschky, P. A. (2024). *The Heterogeneous Productivity Effects of Generative AI*, arXiv:2403.0196.

Li, Z. et al. (2024). "The value, benefits, and concerns of generative ai-powered assistance in writing", in Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, pp. 1–25.

Liu, T. et al. (2025). The illusion of empathy: How ai chatbots shape conversation perception. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 13, pp. 14327-14335).

Noy, S.; Zhang, W. (2023). "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence", in *Science*, vol. 381, no. 6654, pp. 187–192.

Otis, N.G. et al. (2023). *The Uneven Impact of Generative AI on Entrepreneurial Performance*, Tech. Rep. 24-042, Harvard Business School Working Paper.

Peng, S. et al. (2023). *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*, Available at: https://arxiv.org/abs/2302.06590, arXiv:2302.06590.

Qiao, D.; Rui, H.; Xiong, Q. (2024). “AI and freelancers: Has the inflection point arrived?”, in ICIS 2024 Proceedings.

Roldán-Monés, T. 2024. *When GenAI Increases Inequality: Evidence from a University Debating Competition* (London, EsadeEcPol Working Paper).

Schwarcz, D. et al. (2025). *AI-Powered Lawyering: AI Reasoning Models, Retrieval-Augmented Generation, and the Future of Legal Practice*, Minnesota Legal Studies Research Paper No. 25-16 (Minneapolis, MN, University of Minnesota).

Song, F.; Agarwal, A.; Wen, W. 2023. “The Impact of Generative AI on Collaborative Open-Source Software Development: Evidence from GitHub Copilot”, in *SSRN Electronic Journal.*

Suh, D.; Oh, S. (2026). *Generative AI and the Reallocation of Time: Productivity, Leisure, and Fulfilling Work*, Bank of Korea. arXiv:2602.12695.

Teutloff, O. et al. (2024). “Winners and Losers of Generative AI: Early Evidence of Shifts in Freelancer Demand”, in *Journal of Economic Behavior & Organization*, vol. 235, p. 106845.

U.S. Census Bureau. (2025). Business Trends and Outlook Survey. Retrieved from https://www.census.gov/data/experimental-data-products/business-trends-and-outlook-survey.html

Usdan, R.; Kleinberg, J.; Mullainathan, S. (2024). *AI Assistance in General Knowledge Work*.

Yilmaz, E. D., Naumovska, I., & Aggarwal, V. A. (2023). AI-driven labor substitution: Evidence from google translate and ChatGPT.

Yeverechyahu, D.; Mayya, R.; Oestreicher-Singer, G. (2024). *The Impact of Large Language Models on Open-source Innovation: Evidence from GitHub Copilot*, *arXiv:* 2409.08379.

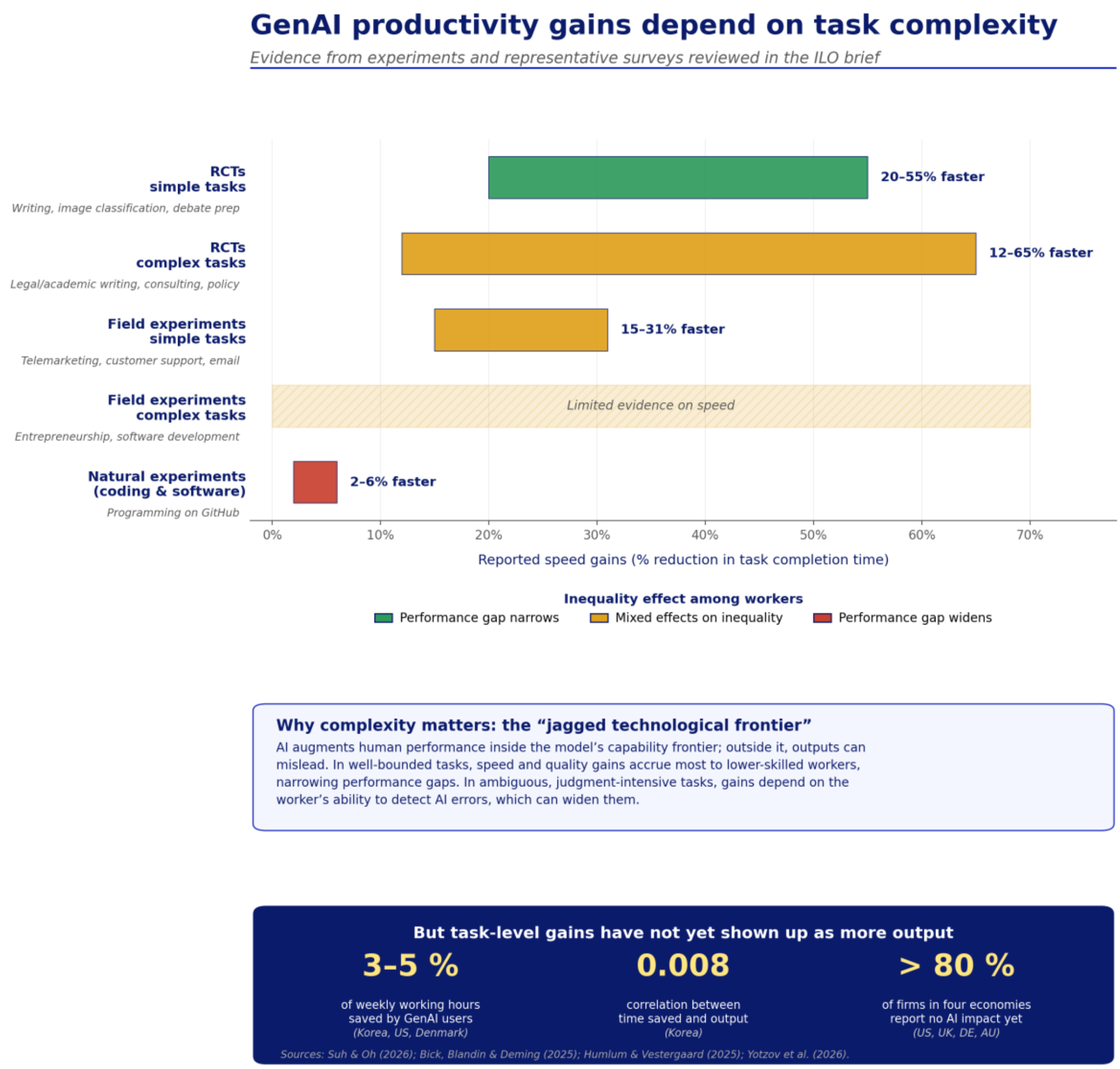
Yotzov, I. et al. (2026). *Firm Data on AI*, NBER Working Paper 34836 (Cambridge, MA, National Bureau of Economic Research).

▶ Table 1: Overview of experiment evidence

Type of Research Design	N	Tasks	Time	Quality	Inequality
RCTs with simple tasks	10	Writing and copy-editing, image classification, debate preparation	Consistently faster (20–55%)	Mixed	Mostly reduction
RCTs with complex tasks	4	Legal/academic writing, consulting, policy assessment	Consistently faster (12–65%)	Mostly positive	Mixed
Field Experiments with simple tasks	3	Telemarketing, customer support, knowledge work	Leans towards faster (15–31%)	Mixed	Mixed
Field Experiments with complex tasks	2	Entrepreneurship, software development	Limited evidence	Mixed	Mixed
Natural Experiments	5	Programming, software development	Mostly positive (2–6%)	Limited evidence	Mostly increase

Caption: Summary of experimental evidence on GenAI and worker productivity from del Rio-Chanona et al. (2025). We classify the direction of findings based on the share of studies reporting on a given outcome: "consistently" indicates more than 80% agreement, "mostly" 70–80%, "leans towards" 60–70%, and "mixed" less than 60%. Studies not reporting on an outcome are excluded from these calculations. "Limited evidence" indicates fewer than half of studies reported on that outcome. Percentages in parentheses show typical effect sizes. Results for Daniotti et al. (2025) were updated from the preprint version reviewed in del Rio-Chanona et al. (2025) to reflect the published findings.

▶ Figure 1: Productivity gains from Generative AI



ILO. *The impact of GenAi on jobs and work organization: a review of empirical evidence,* Research brief, Geneva: International Labour Office, 2026. © ILO. https://doi.org/10.54394/00034628

Contact details
International Labour Organization
Route des Morillons 4
CH-1211 Geneva 22
Switzerland

T: +41 22 799 7239
E: research@ilo.org

DOI : https://doi.org/10.54394/00034628