SAFWAN ALMODHAYAN



OPINION AND ANALYSIS

Editorial date: October 31, 2026

The Two Hours That Returned to No One

> At a US technology company, AI made starting tasks so easy that work leaked into the moments that once stopped it. In Denmark, new tasks appeared while recorded hours and wages remained almost still. The paradox is not that the machine failed to save time, but that task time alone does not decide who receives what it saves.

The Break That Lost Its Right to Be a Break

On February 18, 2026, the Haas School of Business at UC Berkeley published a research interview with Xingqi Maggie Ye. It did not predict the labor market ten years ahead. It concerned eight months of ethnographic fieldwork inside a US technology company with about two hundred employees, conducted with Aruna Ranganathan. Ye observed daily practice and carried out more than forty semistructured interviews, then encountered something she had not assumed in advance: in the observations, making tasks easier with AI did not shrink work but widened its boundaries.1

Ye described three patterns. People began doing work previously assigned to others or not done at all. Sending a request to an AI model became easy enough to do before meetings, over lunch, and in the evening. Multiple workstreams ran at once: an employee reviewed a document as one tool drafted another and a third awaited a decision. Ye said the boundary of what counted as “my work” had expanded. That is an English rendering of the meaning of her published description, not an invented conversation with an employee.

The researchers' account appeared in *Harvard Business Review* on February 9. It reports an ongoing qualitative study, not a randomized experiment or a statistic representing all firms. It does not tell us that everyone using AI worked longer hours. It exposes a precise mechanism: what makes a task cheaper to start can make a new task harder to refuse; what saves minutes in execution can remove the small moments that once separated one piece of work from another.2

Those moments are not holes in the production schedule. Some are the time in which a person thinks before agreeing or recovers attention before deciding. When start-up cost falls, an empty space does not necessarily appear. New demand may appear instead.

Freeing time from a task and freeing a person from a claim on another piece of time are entirely different things.

Three Ledgers, Three Meanings of "We Saved"

If an intelligent assistant prepares a report in half the previous time, the project manager can honestly say that two hours were saved. The finance director may hear that labor cost fell by two hours' pay. The employee may hear a third promise: I will work two hours less.

These are different ledgers; a number cannot move from one to another without evidence. A US study published in *Management Science* in January 2026 used representative surveys of the working population to estimate that respondents' reported time savings from generative AI equaled 1.4 percent of total work hours in a late-2024 wave. This was an estimate from participants' reports, not payroll data or a monetary measure of corporate savings.13

The first is the **task ledger**: time needed to produce a defined outcome at acceptable quality, such as reviewing a contract or summarizing a report, against a clear alternative. The second is the **institutional ledger**: what changed in cash costs, revenue, or capacity to deliver more services. The third is the **human ledger**: what actually changed in hours worked, intensity of attention, after-hours work, and the possibility of reclaiming time for oneself and one's family.

A task can improve on its own ledger while the other two do not move. The human ledger may deteriorate while the institution produces more material. There is no arithmetic contradiction. Contradiction begins when one ledger is presented as all three.

In cash accounting, if an employee's salary remains fixed and no paid hours, contract costs, or hiring need that would otherwise have occurred is reduced, there is no proven cash saving equal to "freed hours × hourly employee cost." Higher quality or other work may create economic value, but that requires independent measurement. Time is not deposited in the corporate treasury like a coin.

This distinction does not diminish the technical achievement. It treats it fairly. A tool that gives a researcher time to read another source or an engineer time to test two alternatives may create real value without removing one riyal from payroll. It is unfair to make every innovation disguise itself as “cost reduction” before its worth can be recognized.

Denmark: Tasks Move While Wages Do Not

On March 13, 2026, Anders Humlum and Emilie Vestergaard released *Still Waters, Rapid Currents: Early Labor Market Transformation under Generative AI*. They linked broad survey evidence with Danish administrative records. The principal late-2024 wave covered about 25,000 workers at 7,000 workplaces in eleven occupations expected to be exposed to conversational models.3

Adoption was associated with worker-reported productivity benefits and new tasks in content writing, oversight of model-output quality, and integration of tools into work. Yet difference-in-differences estimates found no practically meaningful effect on recorded wages or working hours at worker or workplace level during the studied window; confidence intervals excluded average effects larger than about 2 percent under the design.

Did the benefits disappear? No. Was it proved that institutions captured all of them? No. The specific result is that these early benefits did not become a clear material change in examined wages and hours within the measurement horizon. Absorption may happen within the job: new work arises where old work released time.

About 8 percent of tool users at workplaces without institutional initiatives reported new tasks arising from use, compared with about 17 percent where active initiatives existed. These are category comparisons in the study data, not isolated causal estimates of initiative effects. Some new tasks are necessary, such as privacy and quality review. Some may be valuable services previously unaffordable. Some may exist only because producing them became easy.

The researchers did not call every new task waste. This is fundamental: **more work does not itself prove technical failure, just as unchanged pay does not itself prove a fair distribution of gains**.

The study carries strong limits. It was a discussion paper available before final peer review at the cited release. It relies on assumptions in comparisons between adopters and non-adopters, pre-adoption trends, and administrative records’ ability to capture paid hours. Records may miss attention intensity within an hour or evening messages not registered as added work. Two years of a new technology’s diffusion do not settle its effect after a decade.

The evidence nevertheless gives us a better question than “how many minutes did the model save?”: **what work entered the job description after other work left?**

What Did Coal Teach Us, and What Must Not Be Borrowed?

In 1865, the English economist William Stanley Jevons published *The Coal Question*. Greater steam-engine efficiency reduced the coal required for a unit of work but did not necessarily reduce total coal consumption. When useful energy became cheaper, factories, uses, and markets could expand until demand growth exceeded per-unit savings.4

Later literature calls this the rebound effect; its strong form is known as the Jevons paradox. It does not say efficiency always raises consumption. The result depends on demand elasticity, prices, alternatives, new investment, and market size. Rebound may be partial or insufficient to offset savings.

Moving the idea to AI requires a firm boundary around the analogy. A person is not a ton of coal, and markets cannot manufacture more hours in a day. But **demand for tasks** can expand when execution becomes cheaper, even as a person's available time remains fixed.

If reviewing ten strategic hypotheses falls from a week to a day, management may ask for fifty. This may be excellent when the new hypotheses address a valuable problem, or elegant waste when presentation count substitutes for decision quality. Tool speed does not settle it. The marginal value of the new task does: does its benefit exceed the time, resources, and attention it consumes?

Daron Acemoglu and Pascual Restrepo's work on automation and new tasks supplies a deeper labor-economics account. In 2019 they showed how automation may displace human labor from certain tasks while new tasks restore demand. This is a labor-economics framework, not a prediction that every job becomes more exhausting; the result depends on the balance of displacement, reinstatement, productivity, and demand.5

The difference between Jevons and the office is that a market can demand a thousand extra reports, but an employee has one life. Demand elasticity then becomes a question about the limits of work.

The Freed-Hour Ledger

I propose an institutional measurement tool called the **freed-hour ledger**. It is an analytical term proposed here, not an independent academic theory or official accounting standard. Its purpose is to prevent reduced task time from being counted as cash savings or rest before tracing what happened to the minutes afterward.

Suppose recurring work, at a fixed comparison volume and quality, takes S fewer minutes per week. The relation can be simplified as:

$$\Delta H = -S + \Delta N + \Delta V + \Delta O$$

Here ΔH is the change in total minutes actually worked; S is time reduced from comparable tasks; ΔN is time consumed by new tasks; ΔV is additional verification, correction, and integration time; and ΔO is the change in time for other existing tasks. The equation estimates none of these without measurement. It is an illustrative organizational identity, not an empirical model claiming these components explain every change.

In an **explicitly hypothetical example**, a tool removes 120 weekly minutes from an employee's tasks. Fifty minutes move to checking errors and outputs, forty to a new project, and thirty to completing existing work more deeply. Total work time falls by zero. Some outputs may improve or the new project may be worthless; measurement is needed to know which.

In another hypothetical case, 120 minutes are reduced, forty return to other tasks, twenty go to review, and sixty are protected as an earlier departure or less after-hours work. The worker receives an hour not because the model decided so, but because work design allowed it to cross from the task ledger to the human ledger.

Even these accounts require care. Higher task quality may make comparison with yesterday's task unequal; seasonal demand can alter ΔO independently of AI; and verification time must not be counted both within the original task and again as ΔV. A good ledger's first condition is not selling the same minute twice.

When an Employee Sees a Saving the Clock Does Not

In 2025, METR published a randomized experiment with experienced developers working on familiar open-source projects. It included 16 developers and 246 tasks, allowing use of AI tools available from February through June 2025 for some tasks and forbidding them for others.[6]

Before working, participants expected allowed use to reduce time by 24 percent. Afterward, they estimated they had been about 20 percent faster. Measured completion times indicated, in that setting, a 19 percent increase when the tools were allowed.

This is not an argument that AI generally slows programmers. It is a precise example showing that **feeling faster is not measuring speed**. Interacting with a responsive model may feel like progress while reading, reversing, and adapting its suggestions to a complex repository consumes extra minutes.

On February 24, 2026, the same team published an important update. Attempting measurement with newer tools introduced selection bias because some developers no longer wished to work without AI, as well as difficulties timing work with multiple agents. It described newer speed signals as insufficiently reliable for a precise effect estimate, though it believed 2026 tools might improve developer speed more than early-2025 tools.[7]

Management should treat a striking number this way: neither deny nor immortalize it. Ask about skill, task type, tool date, study design, and learning effects. Tomorrow's payroll cannot be derived from today's feeling that work became easier.

The Objection That Sometimes Deserves to Win

Hours can return to their owners. Evidence for this must not be buried because it disturbs a story that work always expands.

In a six-month randomized study covering 66 companies and 7,137 knowledge workers, access to an AI tool embedded in messaging, writing, and meeting applications was assigned randomly. In results published in *American Economic Review: Insights*, actual users—about 80 percent of those given access—spent two fewer hours per week on email during the second half, and their work outside normal hours also declined. Researchers did not find a clear change in task quantity or composition from individual access.[8]

Some space returned to the person. The distinction remains: **two fewer hours on email** are not necessarily **two fewer hours of total work**, even with the positive signal of lower after-hours work. The effect population must also be specified; a user average cannot be assigned to every company employee.

We now have a US institution where researchers observed work intensification, a large Danish sample without material change in recorded hours, and a randomized study where email and after-hours time declined among users. How can all three be true?

They answer different questions at different levels. The first studies the meaning and boundaries of a workday in one institution; the second records wages, hours, and tasks across occupations; the third studies access to a specific tool, digital behavior, and after-hours work in selected companies. The difference among measures is not methodological ornament. It is the heart of the subject.

Sometimes a machine returns time. It has no right to distribute it.

Who Inherits the Freed Hour?

Gary Becker's 1965 paper *A Theory of the Allocation of Time* made time a resource with an opportunity cost. People do not acquire things with money alone; they allocate time among work, consumption, care, and other activities, and an hour's value depends on forgone alternatives.[9]

Applying this inside an institution requires another step. Time a tool frees during paid work does not automatically become an employee's freely owned time; contracts, regulation, and management shape its allocation. Nor can institutions treat the time as free of human cost, because attention, recovery from strain, and family commitments are finite.

Who benefits? Customers may receive a better or faster service. Institutions may gain margin from new work. Employees may gain less repetitive work, rest, or development. Employees may also bear intensified work without higher pay while the institution gains output. Some outcomes can coexist; none follows merely from buying AI licences.

In Gallup results published October 6, 2026, analysis included 7,845 US employees who said they had used AI at work in the previous year. Sixty-three percent reported working faster, while 31 percent said use led their employer to ask for additional responsibilities. These are self-reports, not a causal experiment measuring actual employer requests or added time. They make distribution of gains concrete rather than a mere economic deduction.10

People outside management meetings matter too. If work expands into evening, the cost moves to unrecorded family, study, or care time. If the employee gains flexibility, that is a real benefit even when absent from payroll. Quality of life is not reducible to wages, and welfare figures should not be invented without measurement.

A counter-result deserves celebration. A 2025 *Nature Human Behaviour* study followed an organizational intervention implementing a four-day week without lower pay, with before-and-after data for about 2,896 workers at 141 organizations across six countries and comparison companies. It found improvement in studied measures of burnout, job satisfaction, and health, without the same pattern at comparison firms.11

It was neither an AI study nor a universally randomized experiment of the world's organizations. Its relevance is showing that turning efficiency and organizational capacity into personal time requires an explicit arrangement: redesigning work, defining fewer hours, and committing to a clear result. Rest is not leftover time at the end of a meeting. It is an outcome to choose, protect, and measure.

A Profit Statement That Must Not Be Built on Imaginary Minutes

Labor productivity is generally measured as real output divided by hours worked, not as minutes workers say they saved. That is the definition used within its statistical scope by the US Bureau of Labor Statistics. "Minutes removed from a task" do not establish institutional productivity growth, much less profit growth.12

Imagine an institution with one hundred employees, average hourly labor cost of 150 riyals, and a claim of two hours saved per employee each week across 48 weeks. The spreadsheet produces:

$$100 \times 2 \times 48 \times 150 = 1{,}440{,}000$$

riyals

The calculation is correct. It is not automatically money entering the treasury. If all one hundred employees retain the same pay and ordinary hours, payroll savings of that amount did not occur. If time is reallocated to another service, value is not recorded at the employee's hourly cost; it depends on proven added value, revenue, or avoided cost, after tool, review, and supervision costs.

A simpler, more honest decision criterion is:

Net cash effect = actually avoided costs + realized additional contribution margin − tool, verification, and change costs

This is a management thinking aid, not an official accounting rule or recommendation to book an item. It forbids double-counting. If avoiding an additional hire is counted as a saving, the hypothetical hours of that role cannot also be new revenue. The baseline must be credible: would hiring have occurred without AI; is added demand realizable; is completion quality equivalent?

A manager filling a slide with "saved hours" may think a financial asset has been discovered. Calculations do not become cash because the font grows larger.

If the best effect is freeing an employee to advise a confused customer rather than enter repeated data, that deserves evaluation as better work and customer experience, not an imaginary deduction from her salary.

A Saudi Experiment Whose Results Can Be Falsified

I do not claim the following Saudi experiment exists. I propose a design a bank, service provider, or technology company could use if data, labor, and privacy rules permit, with competent approvals and worker protection.

Choose an activity with a clear quality definition, such as preparing a document package for small-business contracts. Before use, sample tasks and record preparation and review time, output correctness, later corrections, employee departure times, and evening work. Then compare a group using an assistant with one using the ordinary method, through random allocation of units or a justified quasi-experiment.

Do not merely ask, “how much time did you save?” Measure time through samples of work records and voluntary, privacy-protecting diaries, then trace what the organization did with released capacity. Did reviews increase? Did new tasks appear? Did evening work requests fall? Was a genuinely planned external resource avoided? Did customer experience improve after enough time?

Then test a different management decision: one group may expand task volume with demand; another assigns a specified share of time to quality or training; a third protects a break or early departure where service permits. Comparison fails if one group bears harder files or materially different supervision. Service safety and worker rights must be agreed in advance as constraints not traded against minutes.

The freed-hour ledger makes a falsifiable prediction. If shorter tasks automatically reduce work hours, saved minutes will show a stable relation with actual hour reduction without allocation decisions. If destinations differ with work rules and incentives despite similar technology, that supports the proposition that **time distribution is an institutional decision as well as a technical result**. It does not require every company to adopt a shorter week or decline service expansion. It requires choosing knowingly and stating what was measured.

Some institutions’ best return may be higher service quality that cannot be reduced to hourly pay. Another may avoid an expensive contract or reduce paid overtime. A third may give employees time to reclaim life outside work. Each is legitimate if it meets institutional obligations and is not given another outcome’s name.

The Question Before the Next Meeting

The ethnographic study returns us to the moments between tasks: before a meeting, at lunch, in an evening with no clear endpoint. A manager need not formally lengthen the workday, nor an employee request it. A new task only has to become easy, then familiar, then expected.

The tool did not fail to shorten work. It may have succeeded so well that “how long does this take?” became less useful than “why do we do it, and who receives the minutes it leaves behind?”

Time does not create a financial saving merely because it was released on a screen. It does not create human rest merely because a model answered faster. Another decision—perhaps harder than buying the technology—is required: state explicitly what happens to time when compression succeeds.

If an institution tells you tomorrow that AI saved its employees two hours a week, the most useful question may not concern model accuracy or licence count.

Ask: in which ledger did the two hours appear afterward?

References

1. **UC Berkeley Haas School of Business, February 18, 2026.** AI promised to free up workers' time. UC Berkeley Haas researchers found the opposite. **Type:** institutional interview with the original researcher. **Supports:** eight-month ethnography at a roughly 200-person technology firm, more than forty interviews, task expansion, and work entering breaks. **Limit:** one firm and an evolving qualitative study, not a causal national average.
2. **Aruna Ranganathan and Xingqi Maggie Ye, February 9, 2026.** AI Doesn't Reduce Work—It Intensifies It, *Harvard Business Review*. **Type:** research-analysis article by the original researchers presenting an ongoing study; some content may require subscription. **Supports:** self-intensification despite easier task completion. **Limit:** not a peer-reviewed probability-sample paper.
3. **Anders Humlum and Emilie Vestergaard, March 13, 2026.** Still Waters, Rapid Currents: Early Labor Market Transformation under Generative AI, RFBerlin Discussion Paper 078/26. **Type:** preliminary unreviewed working paper at release. **Supports:** about 25,000 workers at 7,000 workplaces in eleven occupations, new tool-related tasks, and no clear material effect on recorded wages and hours, with confidence bounds excluding average effects over about 2 percent. **Limit:** observational comparison with identification assumptions, highly exposed occupations, and an early window.
4. **William Stanley Jevons, 1865.** The Coal Question. **Type:** original historical economics book. **Supports:** greater coal efficiency potentially encouraging total consumption through expanded use. **Limit:** analogy to cognitive-work demand is constrained, not proof of longer hours for every worker.
5. **Daron Acemoglu and Pascual Restrepo, spring 2019.** Automation and New Tasks: How Technology Displaces and Reinstates Labor, *Journal of Economic Perspectives* 33(2):3–30. **Type:** peer-reviewed task-based labor-economics framework. **Supports:** automation displacing some tasks and new tasks reinstating labor demand, without one outcome for every market.
6. **Joel Becker et al., 2025.** Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity, METR. **Type:** 16-developer, 246-task randomized experiment published by the conducting organization. **Supports:** about 19 percent longer measured completion time when early-2025 tools were allowed despite perceived savings. **Limit:** experienced developers, mature projects, and time-specific tools.
7. **METR team, February 24, 2026.** We Are Changing Our Developer Productivity Experiment Design. **Type:** preliminary methodological update by the same team. **Supports:** participant-selection and multi-agent timing problems, and inability to estimate a precise current effect from later results. **Limit:** methodological reflection, not a new controlled experiment proving a specified gain.
8. **Eleanor Wiske Dillon, Sonia Jaffe, Nicole Immorlica, and Christopher Stanton, 2025 study later accepted.** Shifting Work Patterns with Generative AI, *American Economic Review: Insights*; NBER working version. **Type:** randomized field experiment at 66 companies with 7,137 knowledge workers. **Supports:** roughly two fewer weekly email hours among actual users in the second half, and lower after-hours work. **Limit:** email time is not total work time; the study is not reused as this article's opening case or central thesis.
9. **Gary Becker, September 1, 1965.** A Theory of the Allocation of Time, *The Economic Journal* 75(299):493–517. **Type:** foundational peer-reviewed economics research. **Supports:** time as a resource with opportunity cost, not a cash asset created by shortening a task. **Limit:** an economic allocation model does not alone determine contractual working hours.

10. **Gallup, October 6, 2026.** AI Benefits at Work Unevenly Distributed, American Job Quality Study. **Type:** original survey results; 2026 data collected January 26–March 24. **Supports:** 63 percent reporting faster work and 31 percent additional employer-requested responsibilities among 7,845 US employee users. **Limit:** self-reports in a defined population, not causal proof.
11. **Researchers in *Nature Human Behaviour*, online 2025.** Work Time Reduction via a 4-Day Workweek Finds Improvements in Workers' Well-Being, 9:2153–2168. **Type:** peer-reviewed longitudinal organizational intervention with 2,896 workers at 141 organizations across six countries and comparison firms. **Supports:** intentional working-time reduction with improvement in some well-being measures. **Limit:** not an AI study; organizational participation was not universal random assignment.
12. **US Bureau of Labor Statistics, methodological reference.** What Is Labor Productivity?. **Type:** official statistical explanation. **Supports:** labor productivity as real output divided by hours worked, distinct from equating self-reported time savings with cash cost.
13. **Alexander Bick, Adam Blandin, and David Deming, online January 20, 2026.** The Rapid Adoption of Generative AI, *Management Science* 72(10):8440–8456. **Type:** peer-reviewed US survey study. **Supports:** reported time savings equal to 1.4 percent of total work hours in a late-2024 wave. **Limit:** survey estimate, not direct measurement of wages or revenue.