Returns to Intelligence

Suproteem K. Sarkar and Luke Melas-Kyriazi

March 27, 2026

Preliminary and incomplete. Please do not circulate.

Please click here for the latest version

When AI improves, how does usage respond? We document a Jevons-like effect in AI demand—improved capabilities increase usage, and shift AI use toward more complex tasks. We study firms that use the coding platform Cursor. After an increase in model capabilities in late 2025, workers sent 40% more agent messages per week on the platform. The initial increase was larger for lower-complexity messages, though usage shifted over time to higher-complexity messages. The increase was concentrated in smaller, private, and engineering-focused firms, as well as in tasks with more complex dependencies like architecture and database work. The evidence suggests higher latent AI demand for firms with more organizational fluidity, information density, and scalable production. We discuss implications for investment by model providers, and consider how future usage growth relates to both competition between firms and downstream demand.

We thank Mohit Agarwal, Jiafeng Chen, Andrey Fradkin, Spencer Yongwook Kwon, Karl Oskar Schulz, Mirac Suzgun, and Jordan Topoleski for helpful discussions. Sarkar: University of Chicago, Booth School of Business (suproteem@uchicago.edu); Melas-Kyriazi: Cursor (luke@cursor.com).

Global spending on AI infrastructure is expected to approach $1 trillion in 2028 (Patkar et al., 2025). Returns to these investments depend not only on technical progress, but also on how downstream usage responds to improved AI capabilities.

History indicates that technical progress can increase resource use. When steam engines became more coal-efficient, consumption of coal rose rather than fell—cheaper steam power enabled new applications like textile mills, railways, and steamships (Jevons, 1865a; Sorrell, 2009a). Similar effects have been observed for efficiency gains in lighting (Fouquet and Pearson, 2006a). A key driver of these effects is latent downstream demand—in markets with fewer potential new varieties, like vehicles and climate control, more efficient fuel or heating has not increased usage enough to offset the resource savings (Small and Dender, 2007; Aydin et al., 2017a).

How AI progress affects AI usage depends on the adaptability of firms and the latent demand for new applications. Improvements in AI capabilities have reduced the cost of completing some tasks (Erol et al., 2026). If AI usage remains constant—firms do not modify their production processes or find new product applications—such efficiency gains may portend lower AI spending. However, if AI usage increases enough—firms use more AI for existing tasks or find applications to new tasks—AI spending may rise.

This paper studies how AI progress relates to AI usage, and how this relationship varies across firms and tasks. We analyze data on firms that use the coding platform Cursor. We find that capability improvements in late 2025 increased AI usage—workers sent 40% more messages per week on the platform. In addition to using more AI, workers shifted AI requests to target more complex tasks. These responses were concentrated in small, private, and engineering-dense firms, as well as in tasks with more complex dependencies. We discuss how these effects, coupled with the economics of adopting firms, may inform forecasts of AI usage.

Background The recent growth of AI capabilities, revenue, and usage has been driven in large part by the rise of coding agents. Improvements in coding agent capability have been made possible by abundant coding data on the internet and cheap-to-generate verifiable rewards (e.g. unit tests) for reinforcement learning. Coding agents have also become a focus area of AI labs because coding agents accelerate the development of AI itself, speeding up the process of building training infrastructure and running experiments. Some labs believe that these advances could give rise to recursive self-improvement, where the rate of progress accelerates as agents facilitate the research process and conduct research themselves.

The model release cycle of the AI labs creates natural discontinuities in the capabilities of models available to the public. In particular, the recent releases of Claude Opus 4.5 on November 24, 2025 and GPT 5.2 on December 11, 2025 were widely acknowledged as step function increases in model intelligence and usefulness, particularly for coding. Andrej Karpathy, former Director of AI at Tesla, described the impact of these releases: "I don't think I've typed like a line of code probably since December, basically, which is an extremely large change." This sentiment, now common among software engineers, suggests that this moment could be ideal to study how the the usage patterns of firms change in response to improvements in model capabilities.

We study these usage patterns in the context of coding agents by focusing on Cursor, a platform for agent-assisted software development. Cursor is used by firms of varying sizes and sectors, which provides an ideal setting for observing how different firms respond to changes in AI capabilities. We use data on agent interactions, user query classifications, and firm characteristics. The breadth and granularity of these data allow us not only to measure the volume of agent usage before and after major model upgrades, but also to analyze how the scope and complexity of agent interactions evolve.

Increased usage and shift to more complex tasks Following major capability increases in late 2025—including the introduction of Claude Opus 4.5 and GPT-5.2—we document a sharp increase in AI usage: Average weekly requests per worker rose by roughly 40%. This spike began roughly two weeks after the new model releases, as workers returned to work after the holidays, and persisted over the subsequent months. Developers initially used agents to complete more tasks of similar scope and difficulty as before, but over time developers shifted to using agents for more complex tasks, as proxied by the intent of their messages.

Organizational complements predict higher responses We analyze how the increase in AI usage varies across firm and task characteristics, finding that variations in usage are highly heterogeneous.

Our analysis reveals that smaller firms, private companies, and those with a higher density of engineers experienced especially large increases in agent usage. For example, weekly agent messages per worker rose more in the smallest employee tercile than in the largest (44% vs 39%), more in private firms than in public ones (45% vs 37%), and more in the top tercile of firms by engineering density than in the smallest one (44% vs 34%).

Finally, we analyze variation across task categories. The largest increases in agent

usage are found in complex tasks such as architecture design (30%) and DevOps (27%), which require extensive reasoning and integration across systems. There is also significant growth in requests related to learning and understanding existing code (30%), reflecting how coding agents help workers build knowledge in addition to generating code.

Downstream demand and the distribution of surplus In addition to production-side factors like firm size, ownership, and knowledge intensity, responses in AI usage should also depend on market- and demand-side factors.

Market competition may drive up AI usage by creating a technological arms race. For example, if one hedge fund uses more AI to develop trading strategies, rival funds have a stronger incentive to use AI tools. If the size of an adopting market remains fixed, greater demand for AI may increase the margins of model providers without increasing the margins of AI-using firms.

Demand for new varieties may drive up AI usage by creating new applications for the technology. If AI use enables innovation in biology, materials science, and engineering, increased AI capabilities may increase product varieties that use these innovations. If the size of the total market increases, margins may rise for both model providers and AI-using firms.

Related literature This paper relates to the growing literature on demand for frontier models and responses to capability shocks. Demirer et al. (2025) analyze model-usage data and document patterns in pricing, supply, and demand across models and providers. Andrews and Farboodi (2025) study market responses to major model releases, and show that frontier capability improvements have market impact. Our paper uses message-level from professional software developers to study how capability improvements change the level of AI usage and the content and complexity of requests.

A second related literature studies how improvements in AI systems may change production tasks. Jahani et al. (2024) show that some of the gains from better generative models are realized through prompt adaptation: as systems improve, users learn to issue more ambitious and better-targeted prompts. This idea connects to the broader task-based literature, which argues that new technologies reshape the composition of work (Autor et al., 2003; Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2019). Our findings on message complexity fit these perspectives. Our results also complement recent arguments that stronger language models are especially valuable in knowledge-creating settings, where the objective is not only to retrieve or restate known information,

but also to generate new ideas (Cunningham, 2026).

Third, our paper relates to the literature on general-purpose technologies, organizational complements, and heterogeneous adoption across firms (Bresnahan and Trajtenberg, 1995; Brynjolfsson and Hitt, 2000; Bresnahan et al., 2002; Garicano, 2000; Bloom et al., 2012, 2014; Brynjolfsson et al., 2019; Agrawal et al., 2024). Technologies with broad applicability tend to have larger effects in firms that can adapt workflows, redeploy labor, and integrate the technology into core production processes. Our cross-sectional results are consistent with this view.

Fourth, this paper connects to the literature on rebound effects and latent downstream demand. Efficiency improvements do not necessarily reduce resource use; they can instead increase total usage by lowering costs and expanding the range of economically viable applications (Jevons, 1865b; Sorrell, 2009b; Fouquet and Pearson, 2006b; Small and Van Dender, 2007; Aydin et al., 2017b). Our evidence is consistent with a Jevons-like effect in AI demand: better models may raise, rather than reduce, total inference demand by making new applications worthwhile. More broadly, the findings speak to recent discussions of the "marginal returns to intelligence" and the conditions under which capability gains translate into economically meaningful increases in usage (**?**).

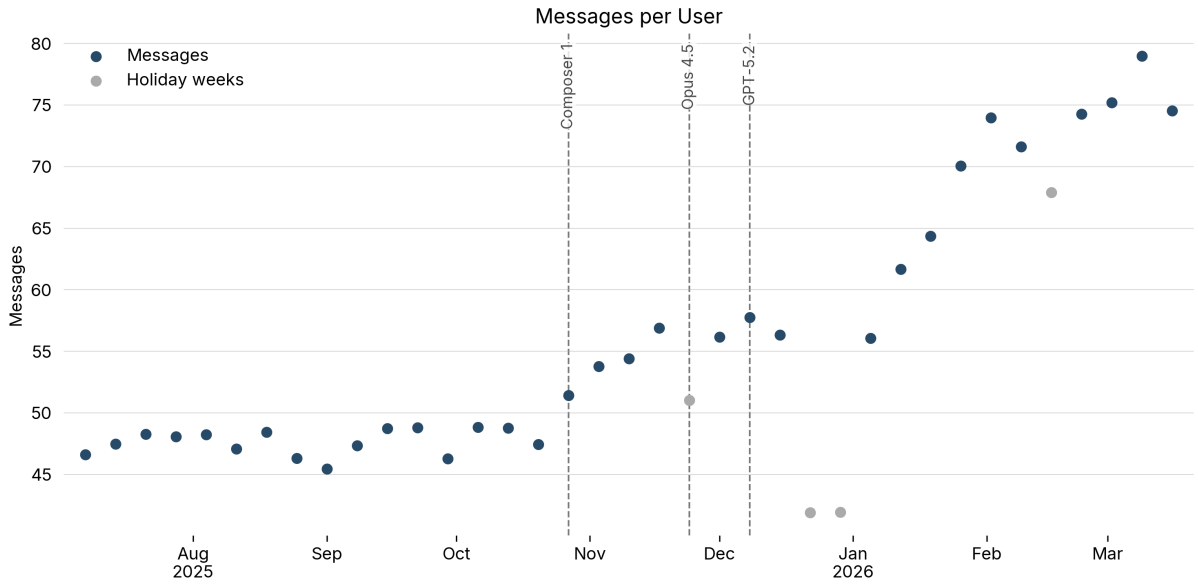
Relative to these literatures, our paper makes three contributions. First, we provide direct evidence on how firms respond to discrete improvements in frontier model capabilities. Second, we show that these capability shocks change not only the level of AI usage, but also the composition of requests, with activity shifting toward more complex tasks. Third, we document heterogeneity across firms and task categories, and link these results to organizational complements and downstream demand.

1. Data

We study firms that use the coding platform Cursor. We use a random sample of 500 firms, uniformly drawn from a subset of firms that joined the platform before July 1, 2025, with at least 50 users, and that used a feature that classifies AI usage on the platform. We also restrict to users that joined before July 1, 2025. We match these firms to total employees, total engineering employees, and publicly traded status using an internal Cursor database. 30.8% of the studied firms are publicly traded. The mean employee count is 9,203, and the median is 1,630.

We link these firms to usage data over time. We analyze data on the number of messages sent per week on the platform between the weeks of July 7, 2025 and March

Figure 1: Usage increased after late-2025 model releases.



Notes – This figure plots the average weekly messages per user on the platform. Vertical lines correspond to three model releases in late 2025. Users send fewer messages in holiday weeks, and data from these weeks are denoted in the figure in gray.

16, 2026. We also analyze data on classifications of messages into complexity, intent, and task category, which is available starting the week of December 15, 2025.

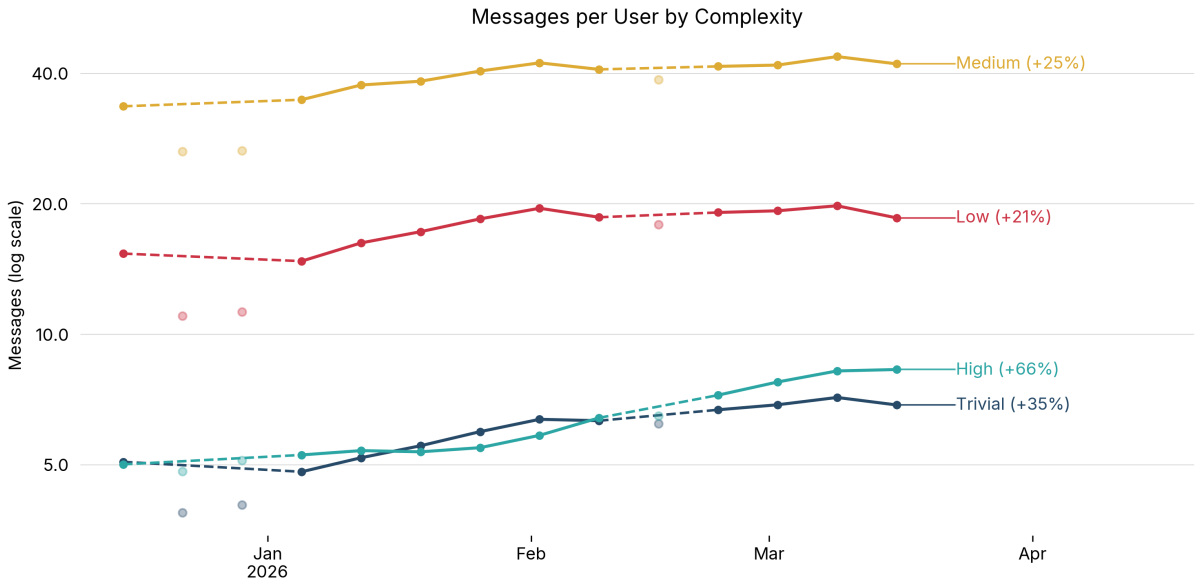
2. Capability Improvements

After improvements to model capabilities in late 2025, AI usage increased and shifted toward more complex tasks. The number of weekly messages sent per user increased by 40%. The largest increase (66%) was in messages classified as high complexity.

2.1. Increased Usage

Figure 1 plots the average number of messages sent over time. The average messages per week rose from 50 in 2025 to 70 in 2026, excluding holiday weeks. On average, this reflected a 40% increase in total messages sent on the platform. This growth in usage suggests that models may have reached a threshold of capability that enabled new use cases or gave users greater trust in model-generated code.

Figure 2: Usage shifted to more complex tasks



Notes – This figure plots the average weekly messages per user, separated by complexity classifications. The vertical axis is placed on a log scale so relative changes are easier to visually compare. Users send fewer messages in holiday weeks, and data from these weeks are denoted in the figure in muted colors, and passed using dashed lines.

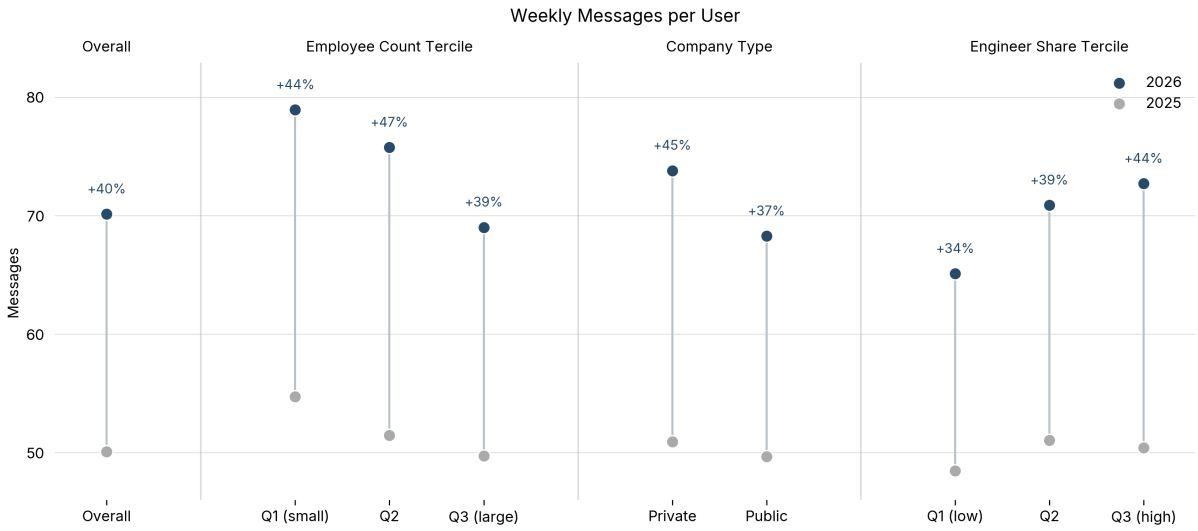
2.2. Shift to More Complex Tasks

Figure 2 plots the average number of messages sent over time, split by complexity classification. Initially, users sent more low- and medium-complexity messages. After 4–6 weeks, the largest increase was in high-complexity messages. Over the entire sample period, the largest increase was in high complexity messages, at 66%. By contrast, low complexity tasks increased by 21% and medium complexity tasks by 25% over the same time period.

Increased model capabilities allow workers to perform new and challenging tasks, but the shift in the distribution of task difficulty lags the general increase in usage. This lag may be partially explained by time it takes users to naturally discover new capabilities over the course of their work, and it may also be explained by users having to reorient their workflows around new model capabilities. For example, in order for developers to fully delegate tasks that last hours or days to their agents, they may have to switch from running agents locally to running them in cloud-hosted environments, which necessarily involves some changes to their workflows.

Another notable finding is that the average number of weekly messages classified

Figure 3: Stronger increases for smaller, private, and engineering-dense firms.



Notes – This figure plots the mean messages per user in 2026 versus 2025, as well as percent increases, across the main sample. The left-most panel corresponds to the overall increase. The second, third, and fourth panels split the sample by firm characteristics. We exclude data from the holiday weeks.

as trivial increased 35%, more than low complexity tasks. This increase may reflect a qualitative change in how people use their coding tools; some developers' primary mode of editing may have switched from editing themselves to conversing with an agent. These developers' first reaction to seeing a small issue in their code may now be to ask the agent to fit it rather than fixing it manually, even when the fix is trivial.

3. Heterogeneous Returns to Intelligence

We find that usage increases are concentrated in smaller, private, and engineering-intensive firms. They are also concentrated in tasks with more complex dependencies.

3.1. Firm Characteristics

Figure 3 plots changes in agent usage across firms of varying size, ownership structure, and engineering density. We find that improvements in model capabilities are heterogeneous across firms, with disproportionate increases in usage across smaller, private, and engineering-dense firms.

First, we compare firms by size, dividing the sample into terciles based on the number

of employees. Smaller and mid-sized firms, which already sent more messages to agents than larger firms on a per-worker basis, saw larger relative increases in usage. Average weekly agent messages per worker rose by 44% in the smallest tercile, 47% in the middle tercile, and 39% in the largest tercile.

Second, we consider ownership structure. Private firms experienced a much larger increase (45% vs 37%) in agent usage than public companies despite already having higher usage before the capability improvements. This difference could reflect faster decision cycles, higher willingness to experiment, sectoral differences, and competitive pressures at private companies.

Third, we examine the share of engineering employees at each firm. Firms with a higher density of engineers saw markedly larger relative increases in agent usage, with a 34% increase for the smallest tercile vs 44% for the largest tercile. This variation could relate to tighter knowledge complementarities, faster adoption, or a greater fraction of work that is directly augmented by coding agents.

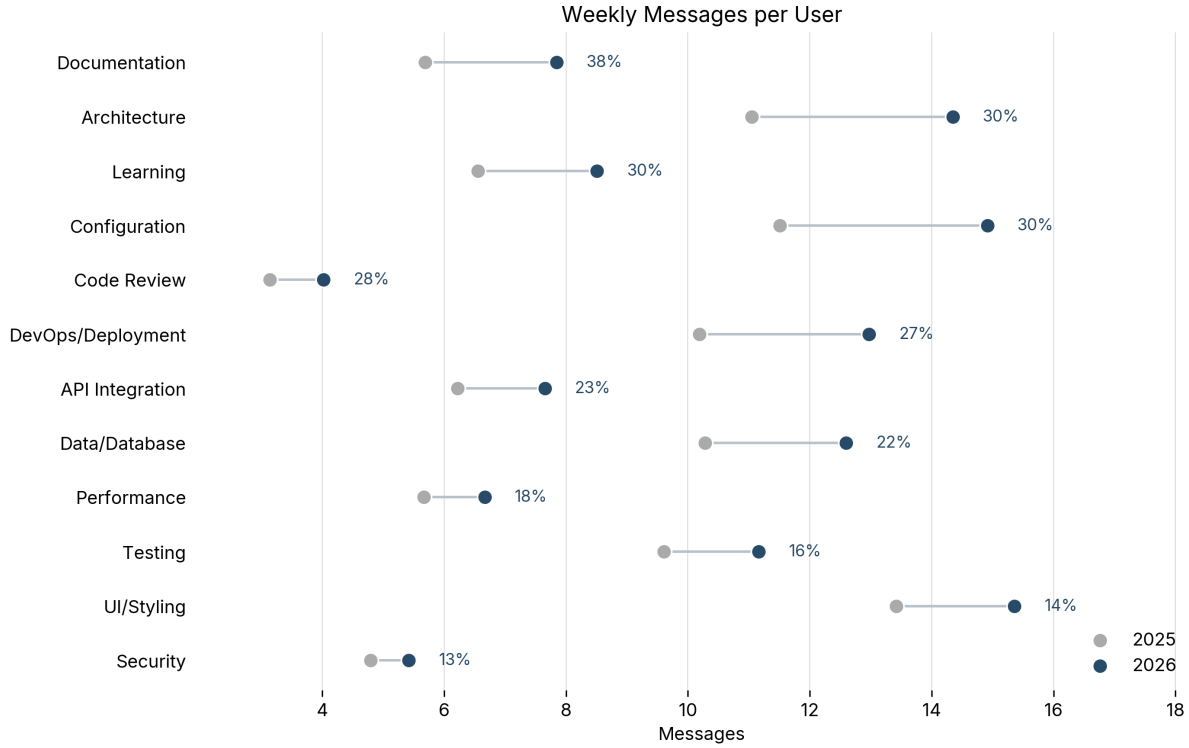
3.2. Task Characteristics

Figure 4 plots changes in agent usage across task categories. The largest increases are seen across tasks that involve reasoning across complex dependencies: architecture (30%) and code review (28%) were among the fastest-growing categories. DevOps and deployment tasks, which can require complex integration work, infrastructure, and production environments, rose by 27%. By contrast, tasks that tend to be simpler and more self-contained tasks such as UI/Styling (14%) saw smaller increases. The growth in tasks involving learning (30%) is also notable; this reflects the degree to which coding agents help workers build knowledge in addition to generating code.

4. Discussion

Our paper finds that increased AI capabilities are associated with higher AI usage. After a capability improvement in late 2025, users of the coding platform Cursor sent more messages and used AI for more complex tasks. These effects were concentrated in smaller, private, and engineering-intensive firms, suggesting the importance of organizational complements. They were also concentrated in tasks with complex dependencies, suggesting that more complex tasks may continue to drive increased AI usage as model capabilities improve.

Figure 4: Stronger increases for tasks with complex dependencies.



Notes – This figure plots the mean messages per user in 2026 versus 2025, as well as percent increases, between the first and final weeks of the conversation classification subsample. The estimated changes are smaller than in the main sample, as this subsample excludes most weeks in 2025. Each row splits the sample by task category classification. We exclude data from the holiday weeks.

Optimal investment by model providers One implication is that the private return to investing in additional model capability may vary considerable across use cases. In consumer-facing chat applications, users may already be at a point where additional intelligence has small effects on the total quantity of use. By contrast, production workflows appear much more capability-bound. Tasks such as architecture design, debugging across large codebases, and database work benefit from models that are reliable over longer horizons and across more complex dependency structures. In these settings, improvements in capability not only make existing usage cheaper, but also expand the set of tasks for which AI is a viable input.

This distinction also suggests that raw model capability and product scaffolding are complements rather than substitutes. Better base models increase the value of tools that

help users apply intelligence in real workflows, such as retrieval over codebases, tool use, memory, and other forms of workflow integration. Conversely, improvements in scaffolding may be especially important in use cases where users are not fully capability-constrained, but are constrained by reliability, coordination costs, or the difficulty of embedding model outputs into production systems. The practical implication is that the optimal investment frontier is unlikely to be one-dimensional. Some downstream demand is enabled by more intelligent models, while some is enabled by integrating existing intelligence into better systems.

Competition and downstream demand A second implication concerns competition and the market-side determinants of AI demand. In some sectors, better models may create an arms-race dynamic. When one firm can use AI to produce, rival firms may face pressure to adopt similar tools to avoid falling behind. This logic is particularly plausible in software and other fast-moving, engineering-intensive markets, where relative performance matters and capability improvements can be translated quickly into product or execution advantages. In such settings, capability gains may raise AI usage even if the direct cost savings per task are modest, because firms are responding to competitive pressure rather static efficiency gains.

In other sectors, however, downstream demand may be much more bounded. If the quantity of final output is limited by market size, regulation, or low product variety, then better AI may primarily reduce the cost of performing existing tasks rather than induce a large increase in usage. Put differently, there are domains in which more intelligence expands the feasible production frontier, and domains in which it mostly lowers the cost of reaching the same frontier. This distinction may help explain why software and code-adjacent work appear particularly responsive to capability shocks: these environments combine high information density, rapid feedback loops, and scope for new product variety. Other occupations may benefit from better models without generating as large increases in total AI usage.

These differences also matter for the distribution of surplus. In highly competitive markets, firms may increase AI usage because relative performance is at stake, yet the gains from that usage may be competed away through lower prices and higher customer expectations. In that case, some of the surplus from better models may accrue to model providers or end users rather than to adopting firms. By contrast, when better AI enables new products, new varieties, or entry into previously infeasible activities, both model providers and adopting firms may benefit.

Taken together, our findings suggest that the economics of AI should be framed not only as a question of substitution, but also as a question of expansion. Better models can increase usage by making AI applicable to tasks that were previously out of reach, and the size of that increase depends on both organizational complements and downstream market structure.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, pp. 1043–1171. 3

_ **and Pascual Restrepo**, "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 2019, *33* (2), 3–30. 3

Agrawal, Ajay, Joshua S. Gans, and Avi Goldfarb, "Artificial intelligence adoption and system-wide change," *Journal of Economics & Management Strategy*, 2024, *33* (2), 327–337. 4

Andrews, Isaiah and Maryam Farboodi, "Do Markets Believe in Transformative AI?," NBER Working Paper 34243, National Bureau of Economic Research September 2025. 3

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333. 3

Aydin, Erdal, Nils Kok, and Dirk Brounen, "Energy Efficiency and Household Behavior: The Rebound Effect in the Residential Sector," *RAND Journal of Economics*, 2017, *48* (3), 749–782. 1

_ **,** _ **, and** _ , "Energy Efficiency and Household Behavior: The Rebound Effect in the Residential Sector," *The RAND Journal of Economics*, 2017, *48* (3), 749–782. 4

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen, "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 2014, *60* (12), 2859–2885. 4

_ **, Raffaella Sadun, and John Van Reenen**, "Americans Do IT Better: US Multinationals and the Productivity Miracle," *American Economic Review*, 2012, *102* (1), 167–201. 4

Bresnahan, Timothy F. and Manuel Trajtenberg, "General Purpose Technologies 'Engines of Growth'?," *Journal of Econometrics*, 1995, *65* (1), 83–108. 4

_ **, Erik Brynjolfsson, and Lorin M. Hitt**, "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *The Quarterly Journal of Economics*, 2002, *117* (1), 339–376. 4

Brynjolfsson, Erik and Lorin M. Hitt, "Beyond Computation: Information Technology, Organizational Transformation and Business Performance," *Journal of Economic Perspectives*, 2000, *14* (4), 23–48. 4

_ **, Daniel Rock, and Chad Syverson**, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, Chicago: University of Chicago Press, 2019, pp. 23–60. 4

Cunningham, Tom, "Knowledge-Creating LLMs," https://tecunningham.github.io/posts/2026-01-29-knowledge-creating-llms.html January 2026. Blog post. 4

Demirer, Mert, Andrey Fradkin, Nadav Tadelis, and Sida Peng, "The Emerging Market for Intelligence: Pricing, Supply, and Demand for LLMs," NBER Working Paper 34608, National Bureau of Economic Research December 2025. 3

Erol, Mehmet Hamza, Batu El, Mirac Suzgun, Mert Yuksekgonul, and James Zou, "Cost-of-Pass: An Economic Framework for Evaluating Language Models," 2026. 1

Fouquet, Roger and Peter J. G. Pearson, "Seven Centuries of Energy Services: The Price and Use of Light in the United Kingdom (1300–2000)," *The Energy Journal*, 2006, *27* (1), 139–178. 1

_ **and** _ , "Seven Centuries of Energy Services: The Price and Use of Light in the United Kingdom (1300–2000)," *The Energy Journal*, 2006, *27* (1), 139–178. 4

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 2000, *108* (5), 874–904. 4

Jahani, Eaman, Benjamin S. Manning, Joe Zhang, Hong-Yi TuYe, Mohammed Alsobay, Christos Nicolaides, Siddharth Suri, and David Holtz, "Prompt Adaptation as a Dynamic Complement in Generative AI Systems," *arXiv preprint arXiv:2407.14333*, 2024. 3

Jevons, William Stanley, *The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-Mines*, London: Macmillan and Co., 1865. 1

_ , *The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-Mines*, London: Macmillan and Co., 1865. 4

Patkar, Vishwas, Lindsay A. Tyler, Carolyn L. Campbell, Vishwanath Tirupattur, Stephen C. Byrd, James Egan, Kyle B. Yellin, and Catherine Liu, "Bridging a $1.5tr Data Center Financing Gap," North America Insight, Morgan Stanley Research July 2025. 1

Small, Kenneth A. and Kurt Van Dender, "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," *The Energy Journal*, 2007, *28* (1), 25–52. 1

_ **and Kurt Van Dender**, "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," *The Energy Journal*, 2007, *28* (1), 25–52. 4

Sorrell, Steve, "Jevons' Paradox Revisited: The Evidence for Backfire from Improved Energy Efficiency," *Energy Policy*, 2009, *37* (4), 1456–1469. 1

_ , "Jevons' Paradox Revisited: The Evidence for Backfire from Improved Energy Efficiency," *Energy Policy*, 2009, *37* (4), 1456–1469. 4