

RESEARCH REPORT

From computing investment to productive adoption

Official surveys measure different stages of AI adoption. Evidence from the United States, Korea, and Japan and a reproducible workflow calculation distinguish computing expenditure, reported use, labor savings, accepted output, and investment returns.

Information cutoff: 2026-10-27

<https://aesenate.org/research/from-computing-investment-to-productive-adoption/>

Computing investment, the adoption of artificial intelligence, and growth in labor productivity are connected economic developments. They are also different observations. A server can be installed without being fully used. A worker can try a generative tool without changing a firm's production process. A task can become faster while verification, integration, or customer acquisition absorbs the time saved. An adequate assessment follows these transitions rather than treating expenditure or reported use as the final economic result.

This report examines official evidence from the United States, Korea, and Japan and sets out a reproducible approach to measuring productive adoption. The sources cover different populations: American businesses and workers, Korean workers and establishments, and Japanese financial institutions. They cannot support a single unadjusted country ranking. Their differences instead help identify the questions that an evaluation must answer about the unit of observation, the intensity of use, and the connection to accepted output.

The central distinction is between access to a technology and a sustained change in production. Investment makes services available; adoption brings those services into a workflow; productive adoption changes the quantity, quality, or resource cost of delivered output. Each transition can fail or take time. Measuring them separately improves both investment appraisal and the interpretation of aggregate statistics.

What the official observations measure

A Federal Reserve research note dated April 3, 2026 compares several American adoption surveys. It reports a firm-weighted AI estimate of 18 percent from the Census Business Trends and Outlook Survey for December 2025, against an employment-weighted estimate of 78 percent from the Survey of Business Uncertainty for November 2025. The note explains that an employment-weighted firm measure approximates employment at adopting firms, rather than the share of individual workers personally using AI. Neither estimate is a productivity gain. [Federal Reserve, monitoring AI adoption.](#)

The Census survey also changed its core question in November 2025, broadening the wording from use in producing goods or services to use in any business function. That break matters when reading the historical series. A rise across the change in wording cannot be assigned entirely to new adoption. The research note records the revised wording and identifies the break; the original wording remains relevant to earlier observations. [Federal Reserve, Figure 2 note and questionnaire appendix](#).

Korea's central bank reports a different kind of evidence. Its November 21, 2025 issue note uses a household survey reflecting the occupational distribution of workers. It reports 51.8 percent using generative AI for work and distinguishes estimated time savings from a potential productivity effect. These are worker-based survey findings and an analytical estimate, not a measured increase in national output per hour. [Bank of Korea, Issue Note 2025-22](#).

A Korea Labor Institute brief published June 24, 2026 provides establishment-based evidence for an earlier period. Using the Workplace Panel Survey, it reports AI adoption of 5.0 percent at the end of 2023. The difference from a later worker survey cannot be interpreted as a directly measured increase of a common adoption rate. The dates, respondents, technology definitions, and survey populations differ. [Korea Labor Institute, Panel Brief 36](#).

For Japan, the Bank of Japan's August 24, 2026 survey report covers 150 financial institutions. It records roughly 80 percent using generative AI and more than 90 percent using or trialing it. The report distinguishes administrative applications from a widening range of operational uses and examines risk management. This is concrete evidence for a financial-sector population, not a national business adoption rate. [Bank of Japan, FY2026 survey](#).

Together, these sources support a narrower but valuable conclusion: AI is reaching different parts of business activity, while the measured extent depends materially on what is counted. A worker trying a tool, a firm authorizing use, and a financial institution running a production application are distinct stages. Comparing the published percentages without those definitions would create an apparent precision unsupported by the evidence.

From capacity to an operating process

The first stage is the supply of computing services. Equipment, buildings, power connections, networks, and software can increase the capacity available to users. Investment expenditure records resources committed during the accounting period. It does not establish that the resulting assets are operating at intended utilization, that customers will pay the expected price, or that downstream users have changed their production methods.

Installed capacity and usable capacity can differ. A facility may have equipment but face a power, cooling, network, or commissioning constraint. A customer may have access to a model but lack permission to use the required data. A firm may purchase licenses while only a small group uses them regularly. The relevant utilization measure should follow the service actually available for economically relevant tasks, rather than infer productive use from the number of units purchased.

The second stage is integration. A useful tool needs a defined task, reliable inputs, a connection to other systems, and a process for exceptions. An answer generated quickly may be unusable if it cannot be linked to a customer record, checked against an authoritative source, or delivered in the format required by the next stage. Integration costs can be incurred in engineering, legal review, employee time, and the redesign of responsibilities.

The third stage is organizational adjustment. A firm may need to change task allocation, staffing, incentives, or the order in which work is performed. A local time saving creates an option, but management still needs to use the released capacity. It can produce more, improve quality, reduce overtime, or allow work to be reassigned. Those outcomes differ economically and should be observed rather than assumed from the technical performance of the tool.

The final stage is sustained output. A successful pilot must continue working across ordinary variation in inputs, staff turnover, customer requirements, and system changes. Its net benefit can decline if model prices rise, errors become more expensive, or the process handles more difficult cases. A credible adoption assessment therefore includes a period of repeated operation and records both typical performance and the costs of exceptional cases.

Counting firms, workers, and uses

The difference between firm and employment weighting can be shown without relying on any disputed adoption estimate. Consider five hypothetical firms. One employs 100 people and four employ ten people each. Suppose the large firm and one small firm have adopted a tool. The firm-weighted adoption rate is two out of five, or 40 percent. The share of employment located at adopting firms is 110 out of 140, or 78.57 percent.

That second percentage does not mean that 78.57 percent of workers use the tool. If only five employees in the large firm and two in the small firm use it, the individual user share is seven out of 140, or 5 percent. All three measures can be correct for the same hypothetical economy. They answer whether firms have adopted, how employment is distributed across adopting firms, and how many workers personally use the technology.

Intensity adds a fourth measure. A person who uses a tool briefly once a month and another who relies on it throughout the day can both answer yes to a broad use question. Frequency, time, task coverage, and the proportion of output affected help distinguish those cases. Even those measures do not guarantee productivity: extensive use may reflect a difficult process, inefficient experimentation, or tasks that would otherwise not have been undertaken.

The Korean worker and establishment sources illustrate the need to preserve the respondent and period when comparing observations. The Japanese financial-institution survey illustrates another boundary: a high adoption rate in a defined regulated sector does not establish broad diffusion to firms outside it. These limitations are not reasons to dismiss the surveys. They are reasons to use each one for the question its design can answer.

A cross-country monitoring table should therefore retain at least the survey date, reference period, respondent, covered population, technology scope, use threshold, and weighting. A column headed adoption rate without

those fields invites a false comparison. If the purpose is to assess diffusion, the best evidence may be a consistent change within each survey rather than the difference in the levels reported by unlike surveys.

A reproducible workflow evaluation

Consider a hypothetical service operation completing 200 accepted cases using 1,000 labor hours in a month. Output is defined as a case meeting the same acceptance standard, not a draft or a generated response. Baseline labor productivity is 200 divided by 1,000, or 0.20 accepted cases per hour. The exercise holds the mix and difficulty of cases constant and treats every hour used in the process consistently.

After adoption, the tool eliminates 180 hours of original task work. Training and workflow adjustment use 40 hours, and additional review uses 60 hours. Total labor input is therefore $1,000 - 180 + 40 + 60 = 920$ hours. If accepted output remains 200 cases, productivity becomes 0.21739 cases per hour, an increase of 8.70 percent. The gross task saving of 18 percent is larger than the net labor saving of 8 percent and differs from the productivity increase.

Hypothetical monthly measure	Baseline	After adoption
Original labor requirement	1,000 hours	1,000 hours
Gross task hours removed	0	180
Additional training and adjustment	0	40
Additional review	0	60
Total labor hours	1,000	920
Accepted cases at unchanged quality	200	200
Accepted cases per labor hour	0.20000	0.21739

The distinction between an hours reduction and a productivity increase is arithmetic. Holding output fixed, an 8 percent fall in hours raises output per hour by 1 divided by 0.92 minus one, or 8.70 percent. Stating the two percentages as though they were identical would introduce a preventable measurement error. For larger changes, the difference becomes more material.

Quality changes the result. If only 190 cases meet the original acceptance standard after adoption, productivity is 190 divided by 920, or 0.20652. The gain over the baseline is then 3.26 percent, not 8.70 percent. If the firm counts all generated drafts as completed cases, it can appear to obtain a much larger gain without delivering more usable output. Acceptance and rework should therefore be specified before the pilot begins.

Costs supply a separate test. Suppose labor costs 30 monetary units per hour and other baseline costs are 2,000. Baseline total cost is 32,000, or 160 per accepted case. After adoption, labor costs 27,600; unchanged other costs remain 2,000; and the new computing service costs 4,000. Total cost becomes 33,600, or 168 per case at unchanged output. Labor productivity has risen while unit cost has increased by 5 percent.

These amounts are illustrative assumptions, not market quotations. They show why a labor-productivity gain and a profitable investment are different claims. A higher service price, greater quality, or future scale could justify the new process, but those benefits need to be demonstrated. Omitting a new purchased input can make a shift from labor to computing look like a free improvement in efficiency.

The counterfactual and the implementation period

A before-and-after comparison is informative only if the underlying workload remains comparable. Demand, case complexity, staff experience, or an unrelated software change can affect performance. If adoption coincides with an easier workload, the observed gain will overstate the technology's contribution. If it coincides with an unusually difficult period, it can understate it. A useful evaluation records these conditions rather than assume that all subsequent changes are attributable to the tool.

Random assignment across suitable comparable tasks or teams can improve causal inference where operationally feasible. It still requires attention to spillovers: employees in a comparison group may learn from treated colleagues, and shared support staff may absorb additional work. A staged rollout can provide another design, but the timing must not simply select the easiest teams first and then attribute their superior results to early adoption.

The evaluation should also preserve the full task boundary. A drafting team may save time while a legal, compliance, or customer-support team handles more exceptions. Measuring only the first team shifts costs outside the recorded process. For a group operating across countries, review can be performed in another subsidiary. The location of the saved hours and the location of the new work should both be identified before a consolidated productivity claim is made.

One-time adjustment and recurrent operation should be reported separately. In the illustrative example, removing the 40 hours of training and adjustment in a later month would reduce total input to 880 hours, provided the gross saving and review requirement remain unchanged. Productivity at 200 accepted cases would then be 0.22727, a 13.64 percent gain over baseline. That is a conditional later-period calculation, not a promise that training or implementation costs disappear.

Conversely, a process can require continuing training as tools, regulations, and customer needs change. Labeling all such costs as one-time can overstate the mature benefit. A defensible business case includes a renewal and maintenance allowance and examines whether the system still works when the original pilot team is no longer providing exceptional attention. Productive adoption is an operating capability rather than a launch event.

Demand and the use of released capacity

Time savings do not automatically become additional output. A firm facing strong demand may use the released capacity to complete more work. Another may reduce overtime or vacancy requirements. A third may retain the same paid hours and improve speed or quality. These choices can affect measured labor productivity, compensation, employment, and profit differently. An evaluation should follow the actual use of capacity rather than equate technical hours saved with jobs eliminated.

Suppose the hypothetical operation still pays for 1,000 hours after adoption but uses the 80 hours released for internal projects that are outside the original case count. Case output per paid hour remains 0.20. That result does not prove there was no benefit; the additional projects may be valuable. Their output must be measured or described separately. Counting the same released hours as both reduced labor input and additional unmeasured work would double-count the benefit.

A capacity gain can also be constrained elsewhere. Faster document preparation does not increase completed transactions if approvals, customer demand, or physical delivery remain the binding constraint. The economic value may lie in shorter queues or better resilience during peaks. Those are legitimate outcomes, but they require measures such as completion time and service reliability alongside output per hour. A single task-speed result cannot establish the productivity of the entire organization.

Pricing and competition influence who receives the gain. A firm can keep cost savings as higher margins, pass them through as lower prices, increase wages, or expand output. Competitive responses may force additional investment and reduce the initial advantage. An adoption study at one firm therefore does not by itself establish the eventual distribution of benefits across producers, workers, customers, and technology suppliers.

This distinction is particularly important when comparing production of computing services with adoption by their customers. Strong demand for chips, data centers, or cloud services can raise output in supplying industries before downstream productivity is visible. The timing can reflect a genuine investment cycle, but it also creates exposure if expected customer benefits fail to support the price and volume of those services. Supplier revenue and user productivity are related evidence, not substitutes.

A case-count target can also change behavior. If staff are assessed only on the number of completed cases, they may allocate the released time to simpler work while complex cases accumulate. The measured average can improve even as service quality deteriorates for customers with difficult requests. The evaluation should therefore retain case complexity, waiting time, and the treatment of unresolved work. A productivity gain is more credible when it survives a stable mix of cases and includes the resources required to complete exceptions.

National accounting and international production

At the firm level, purchasing a computing service can replace labor or other inputs. Labor productivity may rise even if the firm's value added changes little, because purchased services are intermediate costs. At the national level, the supplier's value added may appear elsewhere in the economy or abroad. A complete interpretation needs the production boundary and the location of the supplier, rather than treating the user's gross sales as domestic value added.

A domestically produced capital asset can contribute to current investment and later provide services to users. Imported equipment enters investment spending while imports are deducted in the expenditure accounts. Neither treatment means imported technology cannot support later productivity. It means the immediate contribution to domestic production and the later service contribution are different accounting questions. The same distinction applies when access is purchased as an ongoing foreign service rather than through ownership of an asset.

The ownership of an international technology business also differs from the location of its production. Intellectual property, software development, equipment, and customer support can be located in different economies. Revenue attributed to a corporate group does not establish how much value added occurred in a particular country. Comparing American, Korean, and Japanese adoption with investment therefore requires attention to cross-border purchases and production, rather than matching a national user survey directly to a global firm's sales.

Aggregate productivity can respond through both improvements within firms and changes in market shares. More productive adopters may expand, while less successful firms shrink. An industry average can rise even when many continuing firms show limited improvement. Conversely, early adoption by firms with difficult production problems can weaken a simple correlation between use and current productivity. A causal study needs a design that addresses the selection of adopters and the timing of their decisions.

There are limits to official output measurement as well. Quality improvements can be difficult to observe, especially for services that were previously unavailable. Faster access to advice or better matching can benefit users without a simple increase in a conventional unit count. That difficulty does not justify assigning an arbitrary value to every generated response. It calls for a clear account of what the measured output captures and what remains outside it.

Governance as part of production

Verification and risk management are production inputs where errors carry costs. In financial services, a concise but incorrect summary can trigger rework, a customer loss, or an inappropriate decision. A tool may therefore require human review, access controls, and traceable records. Those requirements reduce the difference between gross technical speed and net resource savings, but they can also make a useful application economically viable by containing its downside.

The Japanese financial-institution evidence is relevant to this stage because it examines use and risk management together. The report's population is explicitly defined, and trials are separated from use. That design is more informative for operational evaluation than combining every expression of interest into a single adoption rate. It still does not establish a causal productivity effect for the surveyed institutions. [Bank of Japan, FY2026 survey scope and usage categories](#).

A productive application needs an escalation route for cases the system cannot handle reliably. The frequency and cost of those cases should enter the evaluation. A pilot built from clean historical examples can understate the burden created by incomplete records, conflicting instructions, or unusual customers. Testing should therefore include representative exceptions and examine the cost of recovering from an error, not simply the share of outputs rated acceptable on first inspection.

Operational dependence is another economic consideration. If a service interruption stops a critical process, a firm may need a fallback workflow whose capacity and maintenance cost reduce the apparent saving. Switching providers can require revalidation, data changes, and employee training. Those costs affect the investment's expected return even where ordinary monthly performance is strong. They belong in the business case alongside subscription prices and hardware expenditure.

Evaluating scale and retention of the benefit

Scaling the illustrative workflow introduces an additional test. A system may perform well on a narrow set of records selected by an experienced team but require more review when extended to other products, languages, or customers. The evaluation should record whether the new cases resemble the pilot and whether additional staff can achieve similar results. Applying the pilot's percentage saving to the entire organization without that evidence treats technical and organizational variation as though it were absent.

A firm should also test retention after the initial period. Changes in source documents, interfaces, customer behavior, and model versions can weaken a previously useful workflow. Monitoring accepted output and exception costs over time helps distinguish a durable process from a temporary demonstration. A sustained saving should be evaluated after ordinary maintenance and oversight rather than on the assumption that the original system can operate indefinitely without further resources.

The cost comparison should include the option to stop or change course. A reversible service purchase can be tested at limited scale, while a specialized facility or extensive internal integration may carry larger exit costs. That difference affects the information needed before commitment. It does not imply that one model is always superior; it means that uncertainty and irreversibility belong in the investment decision alongside expected operating savings.

For a public program, evaluation should additionally distinguish benefits captured by the recipient from benefits available to other firms or users. A subsidized demonstration may produce transferable knowledge, but that claim requires evidence of dissemination and subsequent use. Counting the recipient's expenditure as both a direct benefit and a general productivity spillover would overstate the result. The appropriate evidence follows how a demonstrated process changes behavior beyond the supported project.

Conclusion

The evidence examined here documents important adoption across American businesses, Korean workers and establishments, and Japanese financial institutions. It also demonstrates why published percentages cannot be combined into a common national ranking without harmonizing their populations, timing, and definitions. Evidence of access and use is substantial; the connection to sustained accepted output remains a separate empirical question.

The original workflow calculation shows how that question can be answered at an operational level. Gross task savings, net labor input, quality, purchased computing services, and the use of released capacity must all be recorded. An improvement in labor productivity can coexist with higher unit cost, while valuable service improvements can escape a narrow case count. Following the full sequence provides a sounder basis for computing investment than expenditure growth or an adoption headline alone.

For the aggregate productivity comparison, see [Explaining productivity differences across the Atlantic](#). For the American output and investment accounts, see [The composition of American economic growth](#).

Source addresses

The links below retain the source addresses cited in the analysis. References should be read within their stated periods and scope.

<https://www.federalreserve.gov/econres/notes/feds-notes/monitoring-ai-adoption-in-the-u-s-economy-20260403.html>

<https://www.bok.or.kr/eng/bbs/B0000354/view.do?depth=400409&menuNo=400409&nttId=10094689&programType=newsDataEng&relate=Y>

https://www.kli.re.kr/kli_eng/panelBriefView.es?mid=a20103040000&pblct_sn=10287

<https://www.boj.or.jp/en/research/brp/fsr/data/fsrb260824.pdf>