ANALYSIS.TECH

RESEARCH

GENERATIVE AI MAY IMPROVE EFFICIENCY

ONCE IT IMPROVES WORKERS' EFFECTIVENESS

By Andrew Bartels

Generative AI is the latest shiny new technology that claims to deliver big efficiency benefits through deep automation of business operations. But boosters of every hot new technology – from mainframes to PCs to cloud computing and cell phones – have made similar claims in the past. While generative AI does have the potential to deliver real improvements in business efficiency and productivity, how much of this potential will be realized in practice? I think skepticism is warranted that generative AI will deliver significant increases in productivity and associated big losses in employment.

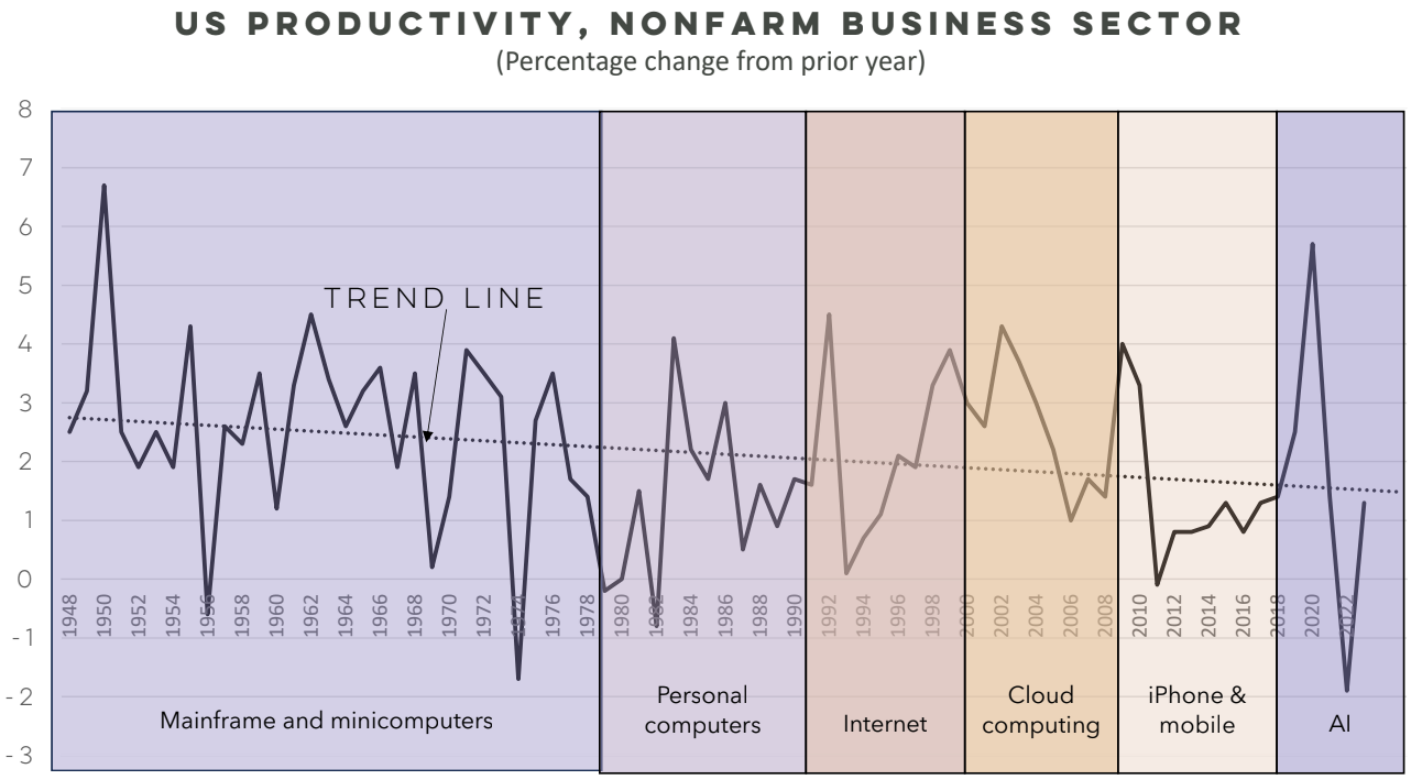
Past experiences with other new technologies suggest that the benefits from generative AI will likely be smaller and slower to be realized than generative AI's advocates think. Meanwhile, there are more mature and proven technologies like low-code, digital process automation (DPA), citizen development, and hundreds of existing process applications that have helped companies achieve automation of processes and slow but steady gains in productivity. The risk is that generative AI investments in isolation will distract companies from devoting resources to the organizational change management required to amplify low-code's and DPA's benefits at scale and improving the performance of existing process apps.

The promise of big improvements in productivity and efficiency has been the major selling proposition behind the massive vendor and investor interests in Chat GPT and the other generative AI products from Microsoft, Google, Amazon, IBM, Meta, and others. Behind these promises are some small and modest success stories. For example, a 2022 MIT study of a trial involving over 400 white-collar professionals found "that the use of ChatGPT increased productivity by 37%, and the time spent on specific tasks decreased by 80%." (The productivity effects of generative AI (ChatGPT) (mit.edu) Wall Street firms are experimenting with using generative AI to prepare the documents that underpin securities offerings, merger and acquisition proposals, and other investment banking activities, replacing high-paid entry-level workers who have done these tasks. Based on these and other trials, IBM CEO Arvind Krishna in 2023 said, "I actually believe that the first set of roles that will get impacted are – what I call – back office, white-collar work." (CNBC.com, IBM CEO says AI will impact white-collar jobs first (cnbc.com)

But it is worth remembering that improving productivity and efficiency is a challenge – certainly at the macroeconomic level, and even at the company level. US productivity (using the US Bureau of Labor Standards' measure for nonfarm labor productivity) did increase in Q4 2023 by 2.6% from the year before. That's an improvement over the 1.9% drop in productivity in all of 2022, and the 1.4% growth in all of 2021. But that improvement is due to a stronger US economy, not to older AI technologies, let alone generative AI, which is barely one year old. Northwestern University economist Robert Gordon,

a long-time skeptic (like me) about technology's improvement of productivity, remains unconvinced that generative AI will have a big impact. Still, other economists are more optimistic, and predict significant improvements in business productivity by the end of this decade (see [Will A.I. Boost Productivity? Companies Sure Hope So. - The New York Times (nytimes.com)]).

Who's right? In my own work as a technology analyst at Forrester Research, I found that information technology investment in this century has produced slow and steady improvements in productivity, but not the quantum leaps suggested by generative AI's boosters. In the 23 years from 2000 to 2023, US labor productivity for the nonfarm business sector increased at an annual rate of just 1.9%, down from an annual rate of 2.2% in the 1990s. These 23 years were the period when cloud computing (SaaS, PaaS, and IaaS), smart phones and mobile apps, social media, robotic process automation, low code, digital process automation and early versions of artificial intelligence and quantum computing transformed the tech industry. There certainly are examples of technology impacting productivity and employment: think of the big drop in toll collector jobs in the US, thanks to RFID and license scanning systems. But overall, the past two decades of information technology innovation and investment have not lifted US productivity growth above the 2.1% annual average since 1948. And the trend line of productivity growth has been steadily slowing in the US.



Source: US Bureau of Economic Analysis

Here are eight reasons I am skeptical that generative AI will have a significant impact on productivity through automating back-office white collar work.

1. Anecdotes about generative AI's benefits are just anecdotes – real evidence of value is still scarce. Each press article and every vendor announcement contain stories of how businesses are using generative AI and AI in general is being put to use to deliver value. In my almost 25 years as a technology analyst, I covered over 300 vendors of software products for commerce, procurement, sourcing, contracts, invoices, and other back-office processes. I have learned to be skeptical of these claims. Vendors routinely exaggerate the benefits of their technology and minimize the challenges of implementation. It's called marketing. In many cases, clients whom I have interviewed did report real benefits from the technology that they adopted. But they often said that the benefits had been overstated, implementation challenges were greater than expected, and the systems had limitations. The newer the technology, the greater was the gap between promise and reality. Older and more mature systems (like low-code and DPA) had easier paths to value, but even here significant organization change was needed to realize the benefits. There is no reason to think that generative AI solutions will be any different.

2. Generative AI still has high potential for errors and hallucinations. There have been many examples of generative AI creating false statements and made-up facts: a lawyer's court filing created with ChatGPT that contained made-up citations; Google's Bard Chatbot claiming the first pictures of planets in other systems came from the James Webb space telescope; ChatGPT, Bing, and Bard all cited the first reference of artificial intelligence in a NY Times article that did not exist. Avoiding these problems requires careful training of AI systems on trustworthy and verified source data and continual human review of the results. Private large learning models (LLM) and Retrieval-Augmented Generation (RAG) tools can also help reduce errors and hallucinations. But all of these necessary steps increase the cost and effort in deriving value from generative AI solutions.

3. Generative AI, like other AIs, still has a steep learning curve. Self-driving cars have been in development for over a dozen years but have still not been deployed beyond carefully controlled tests. The reason? The world is complex and becoming more so with climate change; people's behavior is variable and unpredictable; and unforeseen obstacles like a stopped fire truck or ambulance arise often enough to flummox a self-driving car but not a human driver. Similarly, a generative AI solution that might work in 99% of cases will be stymied by situations that arise 1% of the time. Yet, most business processes and documents inevitably involve many exceptions and detours that even well-trained AI systems won't be able to handle. Will generative AI systems be smart enough – or humble enough – to do that? Humility is not typically a trait that these technology systems have.

4. The benefits of generative AI are much more limited than AI advocates assume. Let's go back to that MIT study cited earlier. Participants in that study were "assigned two occupation-specific and incentivized writing tasks." If writing reports were the only thing that white-collar professionals do, that might be useful. But in fact, writing such reports take up only a small proportion of their time. Indeed, that MIT study went on to report that "Two weeks after the experiment, 33% of the treatment group reported that they had used ChatGPT in their work lives. Respondents who reported not using ChatGPT stated that the technology was not able to integrate or include context-specific knowledge that their occupation necessitates."

5. The financial and environmental costs of generative AI are high and likely to rise for users. While Chat GPT and other generative AI tools have been released as free or open-source tools to use, business-oriented AI solutions will not be free. The NVIDIA and other chips that power generative AI and the data centers that house them cost billions of dollars to build. They also consume massive amounts of electricity – according to a 2023 study, by 2027 AI servers could consume as much electricity in a year as countries like Sweden or the Netherlands do (A.I. Could Soon Need as Much Electricity as an Entire Country - The New York Times (nytimes.com). The producers of generative AI solutions will have to recover those capital and operating costs to make a profit, with resulting high costs to business purchasers of those solutions. Already, Microsoft is charging an additional $30 per user per month just for CoPilot as an add-on to Office 365. A recent MIT research report from January 2024 concluded that the cost-benefit analysis of AI solutions will limit adoption. Looking specifically at AI vision solutions and their costs and benefits, it concluded that "it's only economically sensible to replace human labor with AI in about one-fourth of the jobs where vision is a key component of the work. " (Rethinking AI's impact: MIT CSAIL study reveals economic limits to job automation | MIT CSAIL

6. Generative AI vendors are exhausting high-quality public data and are cutting corners. As the New York Times reported recently, Google, OpenAI, and Meta in 2021 started running out of high-quality, reliable information to train their systems. Meanwhile, providers of existing high-quality and fact-checked information like the NY Times sued generative AI vendors over their unpaid use of these publications' reporting. If these information providers win, AI vendors will find it more expensive and difficult to get good data. Faced with these rising costs, generative AI vendors have already turned to YouTube videos, social media and non-vetted sources. The result could be more and more written reports being created with generative AI solutions whose source data has been polluted with false or misleading information. While private LLM services do avoid some of these risks, they too can struggle if the contributed private data is incomplete or not fully screened.

7. Back-office white-collar work is a small portion of the US workforce. As IBM's CEO indicated, the jobs that are potential targets for replacement by generative AI are back-office work, not customer-facing front-office work where human touch is considered important. These jobs include finance, business analysis, purchasing, human resources, logistics, project management, and legal. Out of a total of 148 million people employment, the 2022 US Occupational Employment Survey showed 9.7 million jobs in business and financial operations, including 3 million financial specialists, 0.9 million in human resources, 1 million in logistics and project management, 0.5 million in purchasing, 0.8 million in management analysis, and 1.1 million in miscellaneous business operations. There were 1.2 million in legal occupations. The combined total of 10.8 million workers in legal and business and financial operations is less than 7% of US jobs in 2022. Small businesses, which tend to be cautious and slow in adopting new technologies, undoubtedly represent 30% to 40% of those jobs. So, even if large companies start adopting generative AI to automate back-office work, the actual potential impact on overall US productivity will be small.

8. Back-office white-collar work is highly collaborative and relies on teams working together. Even in this target area, the actual replacement of workers will be small, because the work is more complex than AI vendors assume. These workers are not simple paper pushers, and the value of their work cannot be evaluated on an individual basis. Except in firms with less than 500 people, these back-office functions operate as teams, with collaborative sharing of ideas, perspectives, knowledge, and best practices. Even at Wall Street firms, the preparation of offering documents is both an educational activity for new investment bankers and a collaboration with older ones. Effectiveness of these teams is thus a multiplicative measure, not an additive one. Replacing one member of a team with an AI bot that can perform 90% of the tasks at 30% of the cost may look like an attractive proposition. But replacing two members of that team with AI bots will result in a loss of 19% of its effectiveness (90% times 90% equals 81%). Replacing three members will create a loss of 27% of its effectiveness, and so on. The more an organization automates these functions through generative AI, the more it reduces the collaboration between people that makes these back-office functions most effective.

While all these factors help explain why the efficiency and productivity benefits of generative AI will be less than proponents claim, the real problem is that generative AI proponents focus on the wrong problem. The real problem: *Back-office white-collar work should be measured on the basis of effectiveness, not efficiency.* If efficiency involves doing work with fewer human resources, effectiveness is measured by making the right decisions at the right time. Back-office white-collar jobs basically involve control, prioritization, and risk mitigation – decisions that impact business results through the balance sheet. Finance positions center on keeping track of revenues and costs, tracking assets and liabilities, reporting financial results accurately and timely, and making sure that capital is invested wisely. HR positions make decisions about which employees get hired, promoted or fired, and with what compensation. Logistics and project managers make sure bottlenecks are anticipated and identified to assure smoothness in operations. Management analysts identify opportunities to improve how the business functions. Lawyers and legal assistants help firms avoid and deal with legal risks and costs. Preparing a merger agreement, a securities underwriting, or bond proposal is more than paperwork – it requires full and accurate presentation of the legal and business risks of a deal. There is a reason that these jobs typically pay $70,000 and more per year: any mistakes that they make can cost the company millions of dollars in damages and lost revenues. Effectiveness, not efficiency, needs to be the measure of success.

So, given these barriers to generative AI solutions, is there a path forward? First, don't let generative AI solutions distract your company from making investments in more mature technologies like digital process automation and low-code that have proven benefits. Second, tiptoe into generative AI by offering line-workers with a portfolio of tools and letting them be the ones to decide whether and where these solutions help them do their jobs more effectively. This bottom-up approach of employee empowerment will not deliver the dramatic improvements in productivity and efficiency that AI advocates promise but it will avoid the billions or even trillions of dollars in bad decisions and lost effectiveness that rapid replacement of back-office white collar workers with AI bots will certainly bring about. Companies that are already pursuing low-code initiatives involve line employees will have an advantage here, because they can leverage the organizational and management processes already in place. This bottom-up adoption of AI tools by workers will not only improve their effectiveness, but it will also slowly and steadily improve their productivity and efficiency.