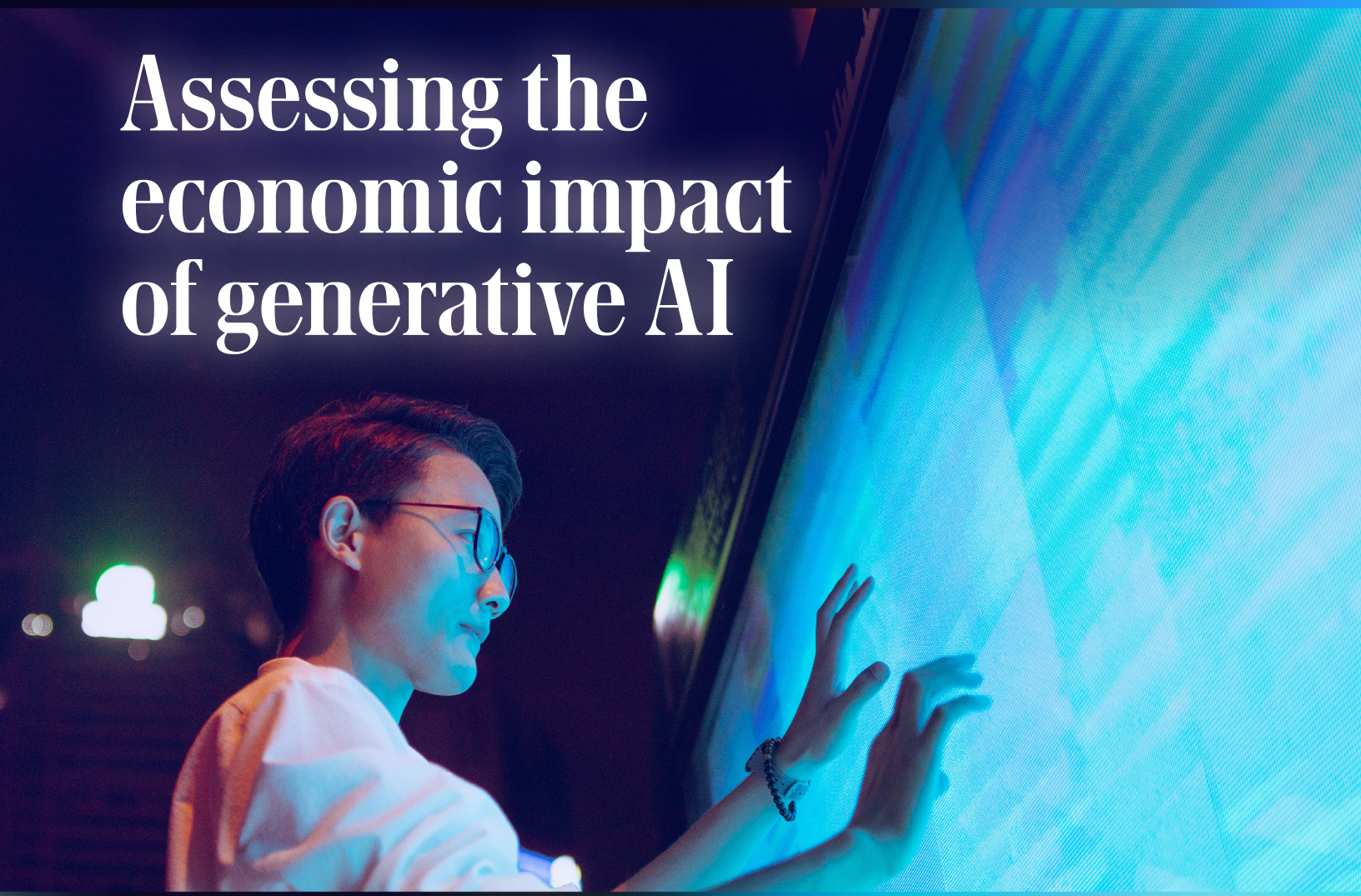
STATE STREET

TRANSLATING PROMISE TO REALITY

Assessing the economic impact of generative AI



Ramu Thiagarajan

Aman Thind

May 2024

Contents

Introduction

The topic of generative artificial intelligence (AI) is top of mind amongst economists, technologists, policy makers and market makers alike. The discussions have been largely optimistic and, in some case, effusive with expectations about transformational and explosive productivity. Are these expectations too optimistic?

Is the advent of generative AI similar to other technological breakthroughs wherein high expectations have been met with deflating reality (e.g., self-driving cars, 3D printing and the metaverse)? Or will it be different this time?

Discussions of generative AI are often, but not always, accompanied by immense optimism about productivity and growth. "Generative AI has the potential to change the world in ways that we can't even imagine," Bill Gates, former CEO of Microsoft has said. "It has the power to create new ideas, products and services that will make our lives easier, more productive and more creative." How can this lofty expectation transform into reality?

There are many hurdles that any technological innovation needs to address before broader adoption that will translate to productivity improvements and growth at scale.

MIT professor Daron Acemoglu opined, "Programs like GPT3 and GPT4 may scuttle a lot of careers but without creating huge productivity improvements on their current path.[1] AI advances are not preordained. The current trajectory is one based on automation. And if that continues, lots of careers will be closed to your children." Will the expected improvement in productivity be accompanied by widespread job losses and structural changes in the labor market? In such a case, widespread optimism around generative AI may be premature.

Generative AI is a global reality and not merely an intellectual curiosity. Interests and expectations around the technology are growing rapidly. According to a 2023 McKinsey Global Survey on the current state of AI, 79 percent of respondents said they have had at least some exposure to the generative AI technology.[2]

The state of the technology and its development continues to advance at a rapid pace as firms are allocating budgets to ongoing research, investment and experimentation.[3]

A survey by Andreessen Horowitz in 2023 shows that almost all companies surveyed planned to increase their spend in the technology from two- to five-fold in 2024, when compared to 2023. As a result, the labor market has also witnessed evolving roles around AI capabilities. For example, according to CNBC, job postings on LinkedIn that mention either AI or generative AI more than doubled globally between July 2021 and July 2023.[4]

The equity market has taken note of these developments. March 2023 marked the genesis of a generative AI-driven stock market rally in 2023, as the release of ChatGPT buoyed enthusiasm for the potential of this transformative technology. The so-called "Magnificent Seven" stocks gained 107 percent in 2023 compared to the overall S&P 500 index at 24 percent. Nvidia, the poster child for generative AI, increased 239 percent in 2023 and continues to advance in 2024.

Despite these tremendous moves in the equity market, corporate capital budgets and hiring trends, mass adoption of generative AI remains in a state of flux. Businesses are exploring pathways for investment and adoption of generative AI in their respective fields while acknowledging the complexities of integration and the technological learning curve. Issues such as bias, security and data privacy, and ethical considerations further perplex their overall decision-making.

In this paper, we examine whether generative AI can indeed herald a new era of economic and technological evolution or fade in similar ways to several previous technological disruptions.

1

To help provide some context, we first discuss the lessons that can be learned from prior historic examples.

2

Next, we outline why we believe the generative AI breakthrough is indeed different and transformative compared to prior technological innovations and outline a framework for assessing its potential to deliver economic gains from a variety of channels.

3

Finally, we posit a number of policy moves that will be critical to help transform the productivity dream of generative AI into the realm of reality.

SECTION 1

Lessons from previous technological disruptions

Challenges in predicting the effects of technological disruptions

Predicting the impact of technological disruptions has proven to be extremely difficult. For example, experts underestimated the transformative effect of automobiles on the economy when they first debuted. Similarly, the advent of the internet was met with both exaggerated short-term expectations (as evidenced by the Dot-Com bubble in 2000-2001) and gross underappreciation of its long-term potential for global connectivity and productivity. Part of the problem is that technological improvements are intangible, difficult to measure and do not show up in productivity statistics in a timely manner. This phenomenon is called the productivity paradox, where the official statistics do not reflect the productivity improvements expected from new technology.[5]

In addition, the effect of technological disruptions is not merely in hours saved but in other dimensions such as increased choices, change of labor demand, and reconfiguration of production process and organizations.[6] For example, use of technology has increased consumer choices in automobiles (*e.g.,* color, features, etc.) while not necessarily increasing productivity.

Some of the effects of technology are reflected in lagged changes in demand. Autor *et. al.* (1997) finds that widespread use of computer technology can explain as much as a 30 to 50 percent increase in the demand for more skilled workers since 1970 with much of demand coming with a lag. As can be seen in these instances, accurately forecasting the trajectory of technological innovations and its broader impact is challenging at best.

Similar to previous technological revolutions, a wave of optimism currently exists around generative AI, accompanied by lofty expectations of productivity growth.[7] Various studies such as Briggs and Kodnani (2023) and McKinsey (2023) suggest that generative AI could raise overall labor productivity growth by around 1.5 percent per annum or deliver total value in the range of US$2.6 trillion to US$4.4 trillion in economic benefits annually when applied across industries. History suggests, however, that many things need to fall in place for the hype of technology to turn into the reality of enhanced productivity.

Lagged effects of technological disruptions

One way to gauge generative AI's predicted impact on productivity is to look at how previous technological innovations have contributed to productivity. History shows that the conversion of technological innovation into tangible productivity gains are often accompanied by a significant lag. For example, it took decades of reconfiguring factory layouts, involving sizeable collateral investments and reengineering work processes, to harness the full productivity improvement potential of the electrification of factories in the early 20th century.[8] Similarly, it was only during the 1995-1999 period that productivity growth saw an increase after the previous two decades of investment in information technology, mainly through the ubiquity of personal computers.[9]

Such examples also show that several structural challenges, such as retraining of workforces and reconfiguration of work processes, need to be addressed to realize the full potential of technological innovation.

For instance, during the mid-1970s and 1980s, the productivity of white-collar information workers actually decreased by 6.6 percent, as employees had to be retrained and readjusted to the use of computers from existing methods and tools.

In addition, as noted in Brynjolfsson and Hitt (1998), companies tend to make a number of changes to their organizational structure to best capitalize on technological innovation, including filling their ranks with highly skilled workers and flattening out their hierarchical structure. These changes eventually help improve efficiency and productivity, but adjusting to them requires time.

These analogies highlight that while the potential for generative AI to enhance productivity is significant, realizing such gains will require not only technological advancements, but also structural and expensive adaptations in overall work processes.

History shows that the conversion of technological innovation into tangible productivity gains are often accompanied by a significant lag.

SECTION 2

Is this time different?

Is generative AI indeed different and unlike anything we have seen before? We believe that this time, it is truly different for two reasons: accessibility and versatility.

History suggests that the most important requirements for the success of disruptive technologies are ease of access and versatility in application. Generative AI excels in both these dimensions. Before we detail each of these attributes, it is important to note that generative AI stands on the shoulders of prior disruptions that have enabled advances that were previously considered impossible. In particular, creating training data sets was not economically viable as recent as a decade ago, as it was prohibitively expensive. The widespread adoption of graphic cards (GPUs), interconnects and frameworks have enabled technology to leverage these high-performance CPU components, allowing huge training data sets (tokens) to be pipelined and processed by the model, thereby making this technology truly transformational.

Lack of access and versatility have resulted in many failures. Apple Newton, for example, a type of personal digital assistants (PDA) with touch-screen capability and portable design, was in many ways a predecessor to the iPhone.

However, its high price and lack of versatility (at its price point) restricted accessibility.[10] Newton was discontinued after five years, and Apple debuted the revolutionary iPhone about a decade later. Google Glass, which included great features like hands-free web navigation, shared the same fate due to its lack of accessibility and high price (US$1,500 in 2013). Thus, for new technology to be adopted with scale and speed, it must possess both accessibility and versatility. A great example of a technology breakthrough that possessed both of these attributes is the smartphone. The successful and rapid adoption of the smartphone was made possible because it was easily accessible to the public and was significantly more versatile than its predecessors.

Accessibility

The landscape of technological innovation has undergone a fundamental shift with the democratization of innovation. Unlike previous technological innovations that were often confined to specialized institutions or companies with substantial resources, generative AI technologies are increasingly accessible to a broad audience including smaller companies and even individuals.

ChatGPT (GPT-3) was a technological breakthrough that became freely accessible to the public. Open-source platforms, cloud computing and collaborative research communities equipped with a wealth of developer tools have lowered the barriers to entry, enabling a diverse pool of talent to contribute to the development and application of generative AI. Community contribution, in particular, has been a key enabler along with readily consumable use cases, such as mobile image augmentation, speech-to-text conversion and translator applications. According to Stanford University's recent (2024) AI Index Report, 65.7 percent of new foundation models released in 2023 were open source, compared to 44.4 percent in 2022 and 33.3 percent in 2021.

The poster child for increased accessibility was the opening of the GPT store by OpenAI in January 2024, through which users can create their own chatbots using generative AI technology and share on the platform. Users can access a wide array of apps tailored to specific needs through this platform, ranging from language learning tools that leverage AI's linguistic capabilities, to business analytics apps that harness AI for data interpretation and decision support.

In addition, platforms like Jasper, Canva and Adobe Firefly allow small and medium-sized enterprises to produce tailor-made and engaging posts or graphical content without the need for expensive resources.[12] Another well-known platform is the Hugging Face Hub, which was introduced in 2022 and became the largest online AI community with over 1.2 million registered users and 10,000 registered organizations. Through such platforms, small startups and individual developers can now leverage powerful generative AI models to create new applications and content, and streamline existing processes.

In fact, the total number of visits to Hugging Face's website has increased from 20 million in March 2023 to 36 million in November 2023.[13] According to OpenAI, more than three million custom GPTs have been created since the announcement of the GPT store.

In addition to the technology, the hardware supporting the technology — namely graphics processing unit (GPU) — has also become more accessible in recent years. Recent advancements in GPU architecture have significantly enhanced their efficiency and computational power. Nvidia's progression from Fermi to Turing and Ampere architectures, alongside AMD's shift from Graphics Core Next (GCN) to RDNA, underscore a leap in performance and energy efficiency. These GPUs now feature expanded core counts and larger memory capacities, far exceeding those from a decade ago. Additionally, the development of enriched software ecosystems, including advanced drivers, development tools and computing frameworks like CUDA and OpenCL, have facilitated broader accessibility of GPUs for general-purpose computing, underscoring their evolving utility beyond traditional graphics applications.

Intel Corporation has recently expanded accessibility by partnering with industry leaders to provide open tools for developers for building and training models without requiring a complete overhaul of their computing system. They have also introduced programs for users, regardless of their technical backgrounds, to familiarize themselves with generative AI usage.[14] Courses on topics like prompt engineering are increasingly being offered by universities, further contributing to the accessibility of the technology. This democratization of generative AI innovation will inevitably accelerate the pace of adoption via a broader range of use cases, thereby making generative AI transformative in its ability to positively impact productivity.

65.7%

Of new foundation models released in 2023 were open source, according to Stanford University's 2024 AI Index Report.

Versatility

Considering generative AI versatility, we are witnessing the power of what is dubbed an "imagenet moment"[15] in natural language processing (NLP). Recent breakthroughs in NLP, powered by foundation models like GPT-4, have unlocked unprecedented capabilities in understanding, generating and interacting with human language.[16] This advancement has enabled immediate and tangible applications across numerous domains, from automating customer service inquiries to enhancing research and development (R&D) through rapid literature review and analysis.[17]

The applications of generative AI in the field of global finance are varied and evolving. We look at a few instances of how such versatility can transform the way the financial industry operates. First, generative AI can streamline processes and improve customer satisfaction through hyper-personalization in communications and generating Requests for Proposals (RFPs), pricing quotes and client onboarding processes. Second, its multi-modal capabilities allow for effective handling of diverse client data, from structured spreadsheets to unstructured pictures and videos, enhancing insights and response times.

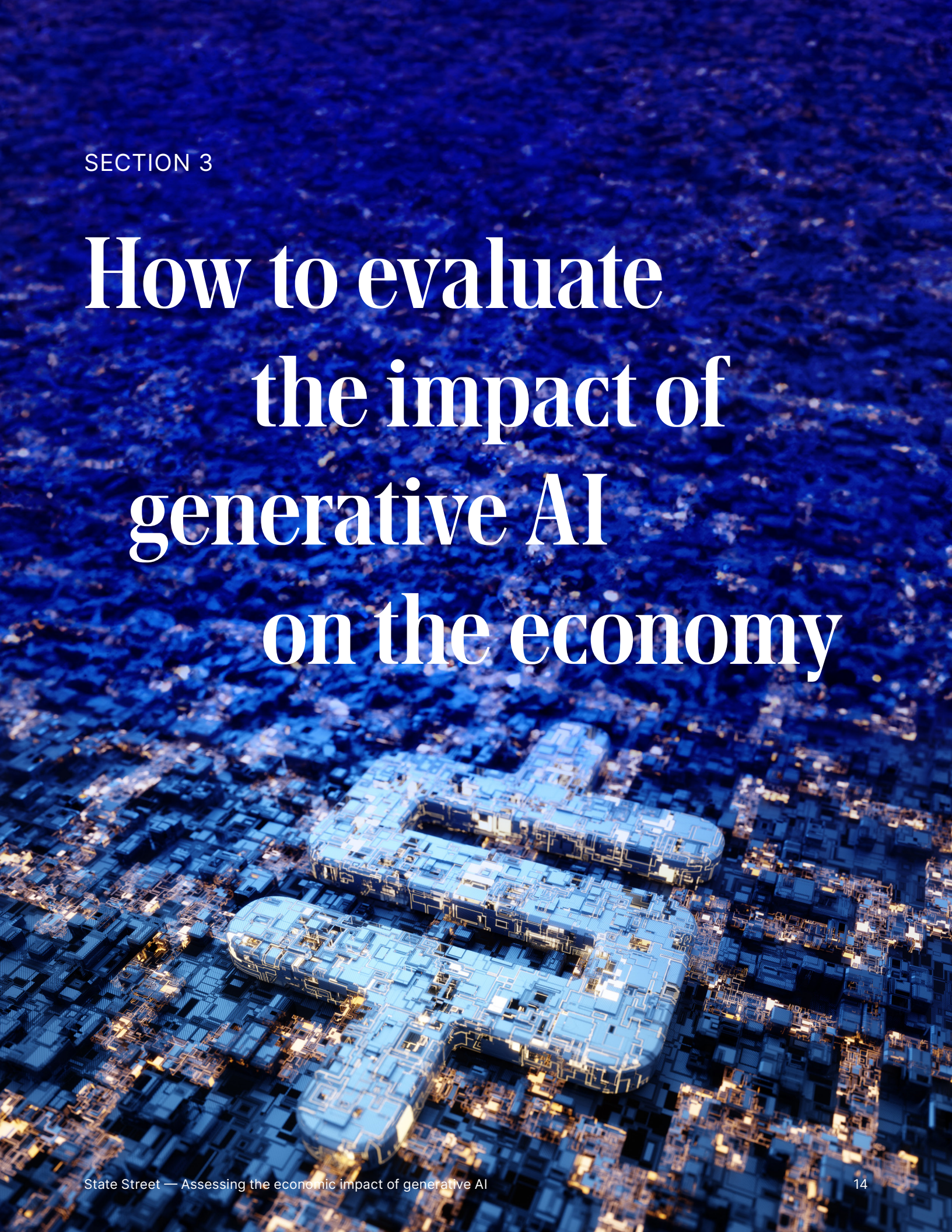
Lastly, its compatibility with low-code or no-code platforms democratizes technology use, allowing quick adaptation to change in the market and fostering innovation.

As can be seen in these examples, the gain is not merely incremental, but transformative in how information is processed, decisions are made and content is created. In other words, generative AI significantly lowers the marginal cost of "knowledge work" through its transformative capability, compared to the internet, which lowers the marginal cost of "knowledge distribution." Thus, the anticipated productivity "J-curve" effect associated with new technology, where initial productivity gains are slow to materialize before exponentially increasing, may be significantly shortened with generative AI.

In summary, we believe generative AI is a technological disruption of a different scale and dimension. It holds great promise as it stands on the shoulders of prior disruptions that have accomplished things previously not thought possible.

SECTION 3

How to evaluate the impact of generative AI on the economy



How does one assess the transformative impact of generative AI on the economy? We have seen isolated examples of the role of generative AI in improving productivity.

For example, a study involving over 700 consultants from the Boston Consulting Group discovered that when AI was used for generating humanlike outputs by conducting analyses, it could enhance worker performance by up to 40 percent.[18] It has been documented that GitHub's generative AI tool, Copilot, has led to a substantial increase in productivity, with up to 46 percent of developers' code across all programming languages generated by GitHub Copilot.[19] The key question, however, remains: How does this translate to the broader economy?

Potential growth of an economy can be thought of as a function of productivity, labor and capital, where higher productivity as well as higher labor or capital generates higher output. Thus, the impact of generative AI on the economy is an aggregation of individual impacts of this disruptive technology on these three factors across different segments of the economy.

Such transformation, however, is not a linear function. In this regard, Acemoglu and Restrepo (2018) and IMF (2024) provide a framework for assessing the pathways of transformations from technological breakthroughs into an economy through different complementary channels:

1. **Labor displacement:** Generative AI adoption can result in a shift of tasks from human labor to AI algorithms, thereby reducing labor income. For instance, according to Hui *et. al.* (2023), freelance writing jobs on an online freelancing platform were down 5 percent after the launch of ChatGPT, and, importantly, earnings were down nearly 10 percent during the same period.
2. **Deepening of automation:** Technological innovation, including generative AI, can increase the productivity of machines in tasks that were already automated, leading to overall productivity gains. For instance, the capability of chatbots can be greatly enhanced through the adoption of generative AI.[20] Techniques such as reinforcement learning through human feedback (RLHF) allow the algorithms to continually improve

efficiently. In addition, generative AI technology is especially adept at finding correlation between data, which allows it to create new automation tasks with ease.

3. **Complementarity:** Adoption of generative AI can increase the value added from occupations with high complementarity between human labor and AI, increasing demand for these occupations. For instance, in industries such as life sciences and chemicals, generative AI has been leveraged for design planning in R&D, significantly reducing research and design time and improving simulation and testing.[21] The technology is also increasingly being used in clinical trials.

   This has resulted in the need for additional skills and jobs for testing and simulation tasks. Generative AI can create synthetic data for training and fine tuning models, both of which have the ability to increase the overall accuracy of traditional AI models.[22] In addition, research by Brynjolfsson, Li and Raymond (2023) found that the introduction of a generative AI tool to guide customer service interactions resulted in a nearly 14 percent increase in productivity. Choi *et. al.* (2023) showed that teams with GPT-4 access significantly improved in efficiency and quality in various legal tasks, including contract drafting.

   These types of deployment of generative AI tools have resulted in the need for additional skills and jobs for testing and simulation tasks.

4. **Creation of new tasks and Industries:** New tasks and industries can be created as a result of broad adoption of new technology. In the case of generative AI, new jobs such as AI ethicists, data curators and algorithm trainers are emerging.[23] Prompt engineering is also a new profession in high demand.[24] The industry continues to grow. According to HAI Artificial Intelligence Index Report 2024, the count of newly-funded AI companies increased to 1,812 in 2023, up 40.6 percent from 2022.

Assessing these effects at the macro level can be extremely difficult due to interaction among these effects and resulting secondary effects. One such effect is what's dubbed "Baumol's disease," where the share of total employment and cost in sectors with high productivity growth decreases while those of low productivity sectors increases. In other words, productivity pays a price as the affected sectors can become a smaller portion of the economy. In the case of generative AI, we feel that it is not likely to be subject to Baumol's disease, partly due to the technology's accessibility and versatility.

Assessing productivity gains at a sector or economy level

As outlined earlier, there are plenty of case studies of economic activity where productivity has been transformed by generative AI. But how does this translate to its impact at the sector or economy level? We believe that this assessment needs a robust framework to make this leap from the micro to the macro level. The work of IMF (2024) provides a meaningful starting point to develop the contours of a framework to assess this impact.

Each economic activity can be viewed and assessed using **two dimensions or lenses**. It is this assessment that will enable one to calculate the impact on productivity from generative AI. The first lens is the exposure of the economic activity to AI. Some activities such as cleaning or ballet dancing don't have "any" AI exposure. If there is no exposure to AI, the chance for improvement is limited. Thus, it is important to assess the exposure of an economic activity to AI. This is the **AI Exposure** axis. The second lens is how complementary is the economic activity to AI and how much is the economic activity likely to be modified, enhanced or supplemented by AI? This is the **AI Complementarity** axis.

AI exposure measures the degree of overlap between AI applications and required human abilities in each occupation.[25]

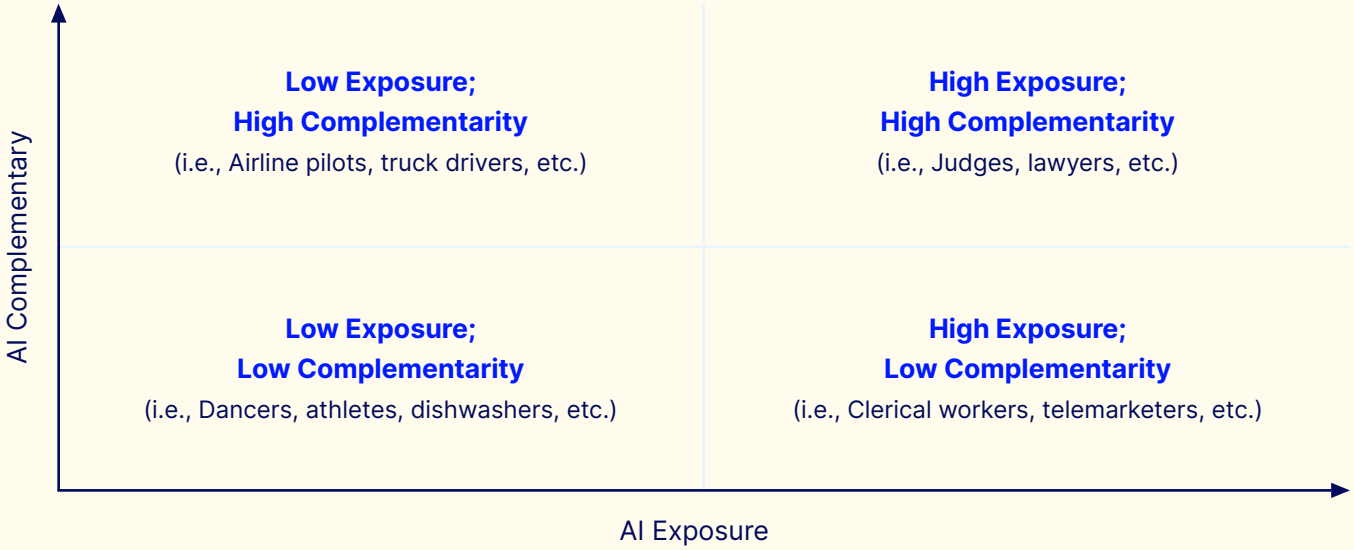
Occupations that are highly exposed to gathering information and analyzing data can greatly benefit from the use of generative AI tools. For example, web developers and judicial law clerks analyze data and process information. Their tasks are highly amenable to automation using generative AI tools. As a result, these occupations are also subject to potential displacement and hence likely to result in job losses at the economy level, but add to productivity enhancement.

AI complementarity measures an occupation's complementarity potential when coupled with the adoption of generative AI. This complementarity potential can be assessed by estimating how much human contribution can be made efficient or improved through the use of technology. The degree of such improvement depends on ethical norm, level of education, required skill levels and physical context of respective occupations.[26] Occupations with high complementarity include judges and clinical researchers as well as airline pilots or truck drivers, which cannot be easily replaced by AI due to ethical norms or physical presence requirements, but can realize large productivity gains. We illustrate these dimensions next.

As seen in Figure 1, economic activities can be categorized into four groups based on the degree of AI exposure and AI complementarity: “high exposure, high complementarity,” “high exposure, low complementarity,” and so on. This framework suggests that certain job families — those involving manual and repetitive tasks and without “shielding” factors such as the level of skills required, social and ethical norms, and physical context — face a heightened risk of obsolescence and job losses but also have great potential for productivity improvement. However, jobs requiring high levels of creativity, complex problem-solving and interpersonal skills, may experience unprecedented productivity boosts without job losses.

For instance, within the asset management industry, the roles of portfolio managers, traders and analysts can be complemented by generative AI technology through using AI-assisted chatbots to ask questions in natural language about their portfolios, trades and counterparty exposures. Software developers can also harness the technology to generate the code, enabling them to focus more on algorithm development.

Figure 1: AI Complementarity versus AI Exposure



Source: IMF (2024)

Full automation or replacement of these occupations, however, will not be likely as certain critical functions such as idea generation, client relationship management, and productionizing the codes require human touch. Therefore, generative AI in the asset management industry complements these occupations through shifting the focus to more creative tasks from manual or repetitive tasks.

Thus, the overall impact of generative AI on an economy is the result of the extant structure of economic activities and how they get categorized in this framework. We explore this next.

Using this framework can greatly help assess the overall impact of generative AI in a given sector or country through the composition of the labor force and the nature of work. In general, countries with a larger share of high-exposure economic activity are likely to experience more disruption relative to those with low exposure. And the impact will be also different depending on the effect of the complementarity dimension.

For example, countries like the United Kingdom and United States, where nearly 70 percent of economic activity are in a high-exposure category, are likely to experience major labor market shifts and a productivity boost from the adoption of generative AI (IMF 2024). Some of these disruptions may already be happening. Juxtapose this shift to a country like India, where the majority of occupations are categorized as low exposure (74 percent)[27] and thus, relatively unlikely to experience the same scale of productivity boost as the UK or the US, though base effects vary.

The UK and the US may face a higher risk of labor displacement for workers in the high-exposure and low-complementarity category. They are also better positioned, however, to reap the benefits of generative AI adoption through a higher share of high-complementarity occupations.

The impact of generative AI on the economy is also a function of how fast and wide the technology can be adopted and how much investment is being made on the technology. According to the AI Readiness Index by Oxford Insights, significant disparity in data, infrastructure, and depth and breadth of the technology sector remains between high- and low-income countries.[28]

The US and UK, for example, rank first and third, respectively, in terms of AI readiness, whereas India ranks 40th. Similar patterns also exist in terms of private investment in the technology. According to the AI Index Report, the US led the private investment in AI technology with US$67.2 billion in 2023, followed by China (US$7.76 billion) and the UK (US$3.78 billion), with India ranking 10th (US$1.39 billion).

US$67.2B

Private investment in AI technology by the US, according to the AI Index Report. The US ranks No. 1 in AI readiness, per the report.

The assessment of productivity gains from the adoption of generative AI on an isolated economic activity is not difficult as there are many metrics that can be used for such evaluation — hours or labor cost saved, improvement in efficiency, etc. Scaling this assessment to a general economic level, however, is not a trivial task. The framework outlined above enables one to aggregate the effects from the broader adoption of generative AI at the level of a sector or economy.

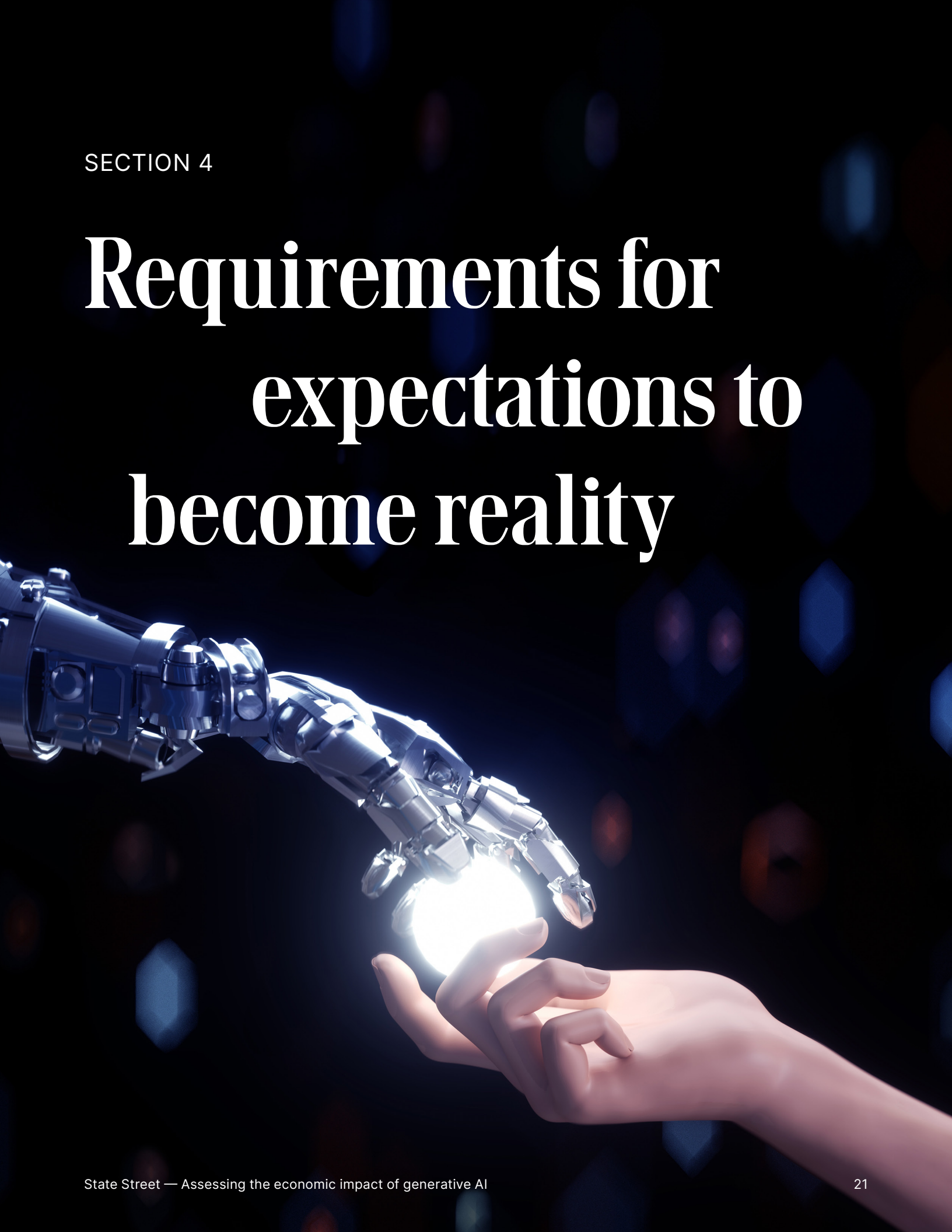
The different projections of potential impact of generative AI help identify the most appropriate policies to prioritize. The goal of policy setting should be to help smooth the transition of activities to harness gains from AI without triggering too much labor market displacement.

Due to accessibility and versatility, it will be easier for companies to adopt generative AI at lower costs than other technological disruptions. And there will be a natural tendency for adoption to be in "high exposure, low complementarity" activities, where adoption will likely be followed by cost saving. As a result, labor displacement in such activities will be the most likely outcome. With no reduction in output, but with reduced input costs, efficiency and productivity will improve. In the long run, however, the ideal technological breakthrough should foster labor growth of a different order to support the ecosystem of improved economic growth. The development of the smartphone, for example, created an entirely new industry of app developers, which resulted in both economic as well as labor growth (albeit of a different nature prior to smartphone development).

Thus, the right policy framework should enable generative AI's role as a tool to unlock new possibilities for innovation and efficiency, leading to potential upskilling and wage growth. Policy should help smooth re-employment challenges and robust transition mechanisms while limiting displacement effects.

SECTION 4

Requirements for expectations to become reality



While we believe that increased accessibility and versatility of generative AI technology can make this time very different, certain requirements need to fall in place in order for expectations to become reality in a non-disruptive manner.

These requirements range across proper policies, regulations and investments to help realize high complementarity benefits of generative AI while help smooth the transition of those with "high exposure, low complementarity" occupations. We share our views on the topic in this section.

Policies

To harness the high complementarity benefits of generative AI while mitigating labor displacement risks, policies must be multidimensional and evidence based, enabling what Professor Acemoglu deemed "human complementarity." Acemoglu notes, "The point of providing data and training is that the algorithm can now do the tasks that humans used to do. That is very different from what I call 'human complementarity' where the algorithm becomes a tool for humans."[29] The key is to help strengthen the role of generative AI for "human complementarity" and limit "human displacement."

Acemoglu and Johnson (2023) propose a few key principles, such as reforming a tax structure to equally favor human labor and technological investment, limiting the use of unproven AI in critical employment decisions, and promoting investment and competition in AI that enhances human expertise. In addition, as noted in Lassebie (2023) and Autor *et. al.* (2003), the government should think carefully about redistributive policies that support worker transition between sectors.

Collectively, these policies should aim to create an ecosystem where generative AI acts as a catalyst for growth and innovation, rather than a source of displacement, with a strong emphasis on equity, accessibility and adaptability to ensure that the benefits of AI are broadly shared across society.

Regulations

Proper regulatory frameworks that balance innovation with safeguarding public interest need to be in place for faster and less disruptive adoption of generative AI technology. Appropriate frameworks are particularly important in areas of competition, inequality, privacy and ethical considerations.[30] In terms of competition, regulations should be in place to prevent the monopolistic control by major tech firms and ensure smaller entities have a fair chance to innovate within the AI space.[31] As suggested by Cremer, de Montjoye and Schweitzer (2019), this involves setting clear definitions around "gatekeeper" platforms and enforceable rules to prevent practices that could unfairly limit competition.

Addressing the exacerbation of inequality across different occupations, income groups and geographies is also an important need in policy setting. In previous broadband internet expansion policies of the early 2000s, these policies often prioritized urban and economically viable areas, leaving rural and low-income communities further behind and amplifying inequalities. Related to generative AI, such policy prioritizations should be thoughtfully considered to mitigate similarly deleterious impacts.

Privacy and ethical considerations are of utmost importance in safeguarding public interest. Privacy regulations akin to the General Data Protection Regulation (GDPR)[32] and the Algorithmic Accountability Act proposed in the US (Congress, 2019) can serve as a model for managing generative AI's implications for personal data and how companies might conduct impact assessments for bias, discrimination and privacy risks in the generative AI development cycle.

Realizing the economic gains from generative AI, thus, requires a balance between technological prowess and thoughtful policy-making that are guided by principles that prioritize human welfare and societal progress.

Conclusion

Today, we stand poised at the door of a potential leap forward in productivity gains through the transformative potential of generative AI. The lessons from past technological disruptions suggest possible failure and importantly a potential lag in realization of productivity gains due to various reasons.

The democratization of innovation, the versatility of the capabilities from this innovation, and the unparalleled advancements in large language models (LLM), however, suggest that this time may be different. While the potential benefits from the adoption of generative AI can be great, the aggregate impact of the technology on the economy is a function of how much economic activities are exposed to and complementary with generative AI. Such an assessment will take time and it is premature to claim broad-based productivity gains for a technology that holds such great promise. Realizing this potential fully depends on our collective ability to craft and implement policies that adhere to the principles of competitive innovation, privacy protection, accountability and human-complementary thinking. The journey ahead is as much about harnessing the power of generative AI as it is about steering its path to one that serves as a catalyst for broad-based prosperity.

Acknowledgements

The authors would like to profusely thank Hanbin Im, Priyaam Roy, and Prashant Parab for their excellent research assistance. Early conversations between Ramu Thiagrajan and Prabhakar Krishnamurthy were helpful in shaping the contours of the paper. The authors would like to also thank Julia McCarthy, Randal Swanberg, Andrew Conn, Michael Metcalfe, Eric Garulay, and Elliot Hentov for productive discussions and comments on earlier versions of the paper.

References

1. Autor, D. H., Levy, F., and Murnane, R. J. (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration." Quarterly Journal or Economics, 118(4), 1279-1333.
2. Acemoglu, D. and Restrepo, P. (2018). "Artificial Intelligence, Automation, and Work." In ***The Economics of Artificial Intelligence: An Agenda***, edited by A. Agrawal, J. Gans, and A. Goldfarb, 197-236. University of Chicago Press.
3. Acemoglu, D. and Johnson, S. (2023). "Power and Progress: Our 1000 Year Struggle over Technology and Prosperity." PubliAffairs
4. Briggs, J., and Kodnani, D. (2023). "The Potentially Large Effects of Artificial Intelligence on Economic Growth." Global Economic Analyst. Goldman Sachs
5. Brynjolffson, E. and McAfee, A. (2014). "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies." W.W. Norton & Company
6. Brnjolffson, E., Li, D., and Raymond, L. (2023). "Generative AI at Work." NBER Working Paper 31161. National Bureau of Economic Research.
7. Cremer, J., de Montjoye, Y., and Schweitzer, H. (2019). "Competition Policy for the Digital Era." European Commission.
8. Felten, E., Raj, M., and Seamans, R. (2021). "Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and Its Potential Uses." Strategic Management Journal 42 (12): 2195-217.
9. Felten, E., Raj, M., and Seamans, R. (2023). "How Will Language Modelers Like ChatGPT Affect Occupations and Industries?" arXiv.org Working Paper.
10. Fiszbein, M., Lafortune, J., Lewis, E., and Tessada, J. (2022). "Powering Up Productivity: The Effects of Electrification on U.S. Manufacturing." NBER Working Paper 28076. National Bureau of Economic Research.
11. Frey, C. B. (2019). "The Technology Trap: Capital, Labor, and Power in the Age of Automation." Princeton University Press.
12. Hui, X., Reshef, O., and Zhou, L. (2023). "The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market." Working Paper 4527336. SSRN.
13. Information Commissioner's Office (ICO). (2021). Data Protection and Privacy Expectations for AI Systems.
14. International Monetary Fund (2024). "Gen-AI: Artificial Intelligence and the Future of Work." IMF Staff Discussion Note SDN/2024/001. International Monetary Fund.
15. Lassebie, J. (2023). "Skill Needs and Policies in the Age of Artificial Intelligence". In: OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market. OECD Publishing.
16. Lo, A. W. and Singh. M. (2023). "From ELIZA to ChatGPT: The Evolution of NLP and Financial Applications." Journal of Portfolio Management, forthcoming.
17. Maslej, N., Fattorini, L., Perrault, R., Parli, V, Reuel, A., Brynjolfsson, E., Etchemendy J., Ligett, K., Lyons, T., Manyika, J., Niebeles, J. C., Shoham, T., Wald, R., and Clark J. (2024). "The AI Index 2024 Annual Report." AI Index Steering Committee, Institute for Human-Centered AI, Stanford University
18. McKinsey & Company (2023). "The Economic Potential of Generative AI: The Next Productivity Frontier".
19. Nestor Maslej, Loredana Fattorini, Raymond Perrault, Vanessa Parli, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, and Jack Clark, "The AI Index 2024 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024.
20. Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., and Li, L. (2023). "Labor Market Exposure to AI: Cross-Country Differences and Distributional Implications." IMF Working Paper 2023/216, International Monetary Fund.
21. United States Congress. (2019). Algorithmic Accountability Act of 2019.
22. Veale, M., Binns, R., and Edwards L. (2018). "Algorithms that Remember: Model Inversion Attacks and Data Protection Law." Philosophical Transactions of the Royal Society A 376(2133).

Endnotes

1. This point is an actively debated topic. For instance, Agrawal, Gans and Goldfarb (2023) disagrees with this notion.
2. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year.
3. Research by Bloomberg Intelligence shows that generative AI could expand to 10-12 percent % of total information technology-related expenditures by 2032 from less than 1 percent today.
4. https://www.cnbc.com/2023/10/04/job-postings-mentioning-ai-more-than-doubled-since-2021-linkedin.html.
5. One explanation for this paradox is that investment in the technology has gone into "unmeasurable sectors" where its productivity effects cannot be measured easily (Griliches). Another explanation proposed by Diewert and Fox is that computers may substitute for other capital such as labor, thus replacing, rather than adding, to some of the productivity gains.
6. We thank Prabhakar Krishnamurthy for this valuable insight.
7. In 2023, venture capital investors invested over US$36 billion into generative AI, which was more than twice as much as in 2022. https://www.economist.com/the-world-ahead/2023/11/13/generative-ai-will-go-mainstream-in-2024?utm_medium=cpc.adword.pd&utm_
8. Research by Fiszbein et al. (2022) shows that electrification led to 1.5 percent higher labor productivity in the US manufacturing.
9. https://cs.stanford.edu/people/eroberts/cs201/projects/productivity-paradox/techbubble.html.
10. Apple Newton went on sale for US$900 in 1993 dollars, which is about US$1,900 today. iPhone, on the other hand, debuted in 2007 at US$499, which is around US$733 in 2024.
11. GPT-4 is currently available to public at US$20/month.
12. GenAI Can Help Small Companies Level the Playing Field (hbr.org).
13. https://originality.ai/blog/huggingface-statistics#:~:text=How%20many%20people%20use%20Hugging,of%20June%202022%20(Source).&text=As%20of%202022%2C%20Hugging%20Face,AI%20community%20online%20(Source).
14. How AI is accessible to all, regardless of technical background (washingtonpost.com).
15. This term was used in Lo and Singh (2023), which refers to the development of generic models that can be used for multiple tasks.
16. The development of foundation model was enabled through the transformer attention mechanism, which is the ability to train models on trillions of tokens with the help of innovation in HPC, GPU, and DPU (data processing units). Computing power like this was previously reserved only for military and scientific research.
17. Generative AI is also capable of generating visual contents such as images and videos. Users can enter a textual prompt describing what type of image or video they want, and the generative AI tool will process the input to generate the contents.

18. https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-can-boost-highly-skilled-workers-productivity.
19. https://github.blog/2023-05-09-how-companies-are-boosting-productivity-with-generative-ai/#:~:text=Generative%20AI%20can%20boost%20developer,can%20inform%20better%20decision%2Dmaking.
20. https://www.kellton.com/kellton-tech-blog/generative-ai-chatbots-gamechanger-or-doomslayer-to-intelligent-conversations.
21. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier.
22. The Amazing Ways Snowflake Uses Generative AI For Synthetic Data (forbes.com).
23. https://www.forbes.com/sites/bernardmarr/2023/12/27/how-generative-ai-will-change-all-of-our-jobs-in-2024/?sh=6f63777174ae.
24. The New 'AI Psychologists': The Rise of Prompt Engineers (forbes.com).
25. Felten, Raj, and Seamans (2021, 2023).
26. Pizzinelli et al. (2023).
27. IMF (2024).
28. Pp 7 and 8, https://oxfordinsights.com/wp-content/uploads/2023/12/2023-Government-AI-Readiness-Index-2.pdf.
29. Goldman Sachs (2023) estimates that generative AI could affect approximately 300 million jobs worldwide over the next decade.
30. Frey (2019) points out the necessity for policies that balance the acceleration of technological adoption with the well-being of the workforce.
31. Some efforts are already underway, such as the European Union's Digital Market Act or the EU Commission's call for insights from all stakeholders on the level of competition in virtual worlds and generative AI.
32. As noted by Veale, Binns, and Edwards (2018), the EU GDPR does not cover models explicitly. Thus, privacy regulations for generative AI should include clear guidelines on consent for AI-generated insights and the use of synthetic data to train AI models (ICO, 2021).

State Street Corporation
One Congress Street, Boston, MA 02114-2016

www.statestreet.com

6647443.1.1.GBL
Expiration date: May 31, 2026