AI and Productivity in Canada

Hans Martinez

2025-08-26

We study the impact of generative AI on productivity in Canada. Using production functions and firm-level data on Canadian firms, we aim to measure if there had been any productivity gains from the adoption of generative AI. The main challenge of identifying a causal effect of generative AI on productivity is that more productive firms may be more likely to adopt AI technologies. By modeling the firms' decision to adopt AI and leveraging from linking different datasets, we propose a clear identification strategy.

Introduction

The large language model (LLM) ChatGPT debuted in November 2022, marking a significant milestone in the development of generative AI technologies. Since its introduction, generative AI has sparked widespread interest and debate regarding its potential to increase productivity at the individual and firm level.

What is the evidence of generative AI increasing productivity? A growing literature points to heterogeneous effects. In most cases, low-skilled workers seem to be the most benefited, but high-skilled workers could even decrease output or quality of their work. There is also evidence that AI could harm productivity. At the firm level, usage of AI —broadly defined— has been linked to higher productivity, but the literature is still scarce.

Low-skilled and less-experienced customer assistants improved significantly across different measures of productivity, after the introduction of a generative-AI based conversational assistant (Brynjolfsson et al. 2023). Access to the AI tool helped newer agents move across the experience curve more quickly. In contrast, there were minimal impacts on the productivity of more-skilled more-experienced agents. Importantly, the authors find evidence that AI assistance may decrease the quality of conversations by the most skilled agents.

In another experiment conducted at Ant Group, a large Chinese big tech company, researchers found similar results (Gambacorta et al. 2024). Following the introduction of CodeFuse, a large language model (LLM) designed to assist software programmer teams in coding, productivity increased by more than 50%. Again, the higher gains in productivity were statistically significant only among entry-level and junior staff. When looking into why senior employees had a lower increase in productivity, the authors find that the senior employees used the gen AI less frequently.

For more complex tasks, such as code development, generative AI might reduce the output and quality of less-experienced users. Using observational data from 36,000 GitHub accounts, Kreitmeir and Raschky (2024) documented that Italy's ban on ChatGPT increased output quantity and quality for less experienced users. Experienced users showed a decrease in more routine tasks. The authors argue that less-experienced users might take longer to identify and debug generative AI's faulty code.

Both of these effects —higher productivity gains for low-skilled workers and productivity decreases for complex tasks— are observed in the experiment conducted at Boston Consulting Group, a global management consulting firm (Dell'Acqua et al. 2023). In this experiment, researchers assigned consultants at random to three different conditions: no AI access, GPT-4 AI access, and GPT-4 AI access with prompt engineering overview. The authors find that consultants using AI were significantly more productive and produced higher quality results. Importantly, consultants below the average performance threshold increase by 43% and those above increasing by 17% compared to their owns scores. Furthermore, for complex tasks that the researchers considered to be outside the AI's current capabilities, consultants using AI were 19 percentage points less likely to produce correct solutions compared to those without AI.

At the individual level, these studies highlight two key points: First, AI expertise matters. General-purpose AI, when used for specialized tasks like coding or tasks beyond its capabilities, can reduce productivity by increasing error detection and correction time, as seen in the ChatGPT ban in Italy and the Boston Consulting Group experiment. Second, specialized AI significantly boosts productivity for low-skilled, less-experienced workers. This makes sense if the AI is trained on data generated by the top performers. This may be the case of the customer assistants. The AI could have been suggesting solutions the high skilled workers would take.

At the firm level, Czarnitzki, Fernández, and Rammer (2023) finds a positive and significant association between the use of AI and firm productivity. The authors use a production function approach and survey data on German firms from 2019. Limited to a cross-section of German firms and a very small panel, the authors cannot use popular methods of production function estimation. Consequently, their study is limited to instrumental variables (IV) methods to account for the possible endogeneity arising from more productive firms being the ones investing more in AI. They employ as instruments —variables correlated with AI use but uncorrelated with productivity— the sector-level AI use, past innovation expenditure, and resistance to innovation by the firms' workers. Their data allows the researchers to measure AI usage with indicators (e.g. whether they use or not AI) and AI usage intensity (e.g. whether firms use AI in products/services, automation of processes, communication with customers, etc.). The authors define AI broadly because the 2019 survey lists only a few AI methods such as natural language understanding, image recognition, machine learning, and knowledge-based systems.

Finally, I could not find evidence or research work linking Canadian firms or workers with generative AI and productivity. However, Canada might be an interesting country in which to study the link between productivity and generative AI. This is due to Canada's diverse industry sectors, including a non-trivial professional, scientific and technical services. Additionally, understanding the impact of generative AI in Canada could provide valuable insights for policymakers and businesses aiming to close the productivity gap with other advanced economies.

Open questions

Even though there is a growing interest in link between generative AI and productivity, there are still many open questions. In addition, there could be an area of opportunity for application in the Canadian market as I could not find evidence for Canadian firms or workers.

Among the possible research questions:

- What is the evidence of the increase in labor productivity due to AI implementation in Canadian services firms?
- What are the implementation strategies Canadian firms are taking to implement the use of AI by their employees?
- What are the challenges firms are facing to have their employees use AI to increase their productivity?
- What have been the unexpected negative effects of implementing AI technologies in the workplace?
- What are the common uses of AI by the employees?
- What are the commonly missed opportunities of using AI in the workplace?
- What are the most common task employees use to automatize with AI?
- Had there been any significant differences between early vs late AI adopters in the workplace?
- Are there any efforts from firms to level up late adopters?
- Are firms considering the heterogeneous effects on productivity by generative AI?
- In the long-term, will low-skilled workers develop their level of expertise, or will they keep relying on AI and hinder their expertise development?

Identification Strategies

The problem is we can't just include a AI_{jt} dummy variable into a production function regression:

$$y_{jt} = g(k_{jt}, l_{jt}, m_{jt}) + AI_{jt} + \omega_{jt} + \varepsilon_{jt}$$

because it is likely that $\mathbb{E}[\omega_{jt}|AI] \neq 0$. In other words, more productive firms are more likely to choose to use AI.

Low case variables are the logs of the levels $log(Y_{jt}) = y_{jt}$.

Here ω_{jt} represents the persisten part of productivity and ε_{jt} is the measurement error (or output shocks). It is usually assume that the persistent part of the error term follows a markov process, in particular an AR(1),

$$\omega_{jt} = h(\omega_{jt-1}) + \eta_{jt}$$

Approach 1: Are firms using AI more productive?

One strategy is to do what we do with other indicator variables that are endogenous to productivity, like being an exporter.

In those cases, we exclude the dummy variable from the production function regression,

$$y_{jt} = g(k_{jt}, l_{jt}, m_{jt}) + \omega_{jt} + \varepsilon_{jt}$$

We can estimate the production function using GNR(2020), or other method, and productivity. With estimates of firm-level productivity, we can compare the productivity of firms choosing and not choosing to use AI. Of course, here we could not claim a causal effect. This will only serve to test our assumption that AI adopting firms are indeed more likely to display higher productivity.

Strategy II

We can think about the model a little bit more. The intuition of the identification in GNR(2020) is that firms choose intermediates after they observe their productivity shock. Because intermediates are flexible (they choose the quantity they want in the current period), then we again have that $\mathbb{E}[\omega_{jt}|M_{jt}] \neq 0$.

However, starting from the firm's cost-minimization problem,

$$\max_{\substack{K_{jt}, L_{jt}, M_{jt} \\ AI_{jt}}} \pi = P_t \mathbb{E}[Y_{jt}] - r_t K_{jt} - w_t L_{jt} - c_t M_{jt}$$

$$\text{s.t.} Y_{jt} = G(K_{jt}, L_{jt}, M_{jt}) \exp^{\omega_{jt} + \varepsilon_{jt}}$$

Taking the FOC wrt M_{jt}, multiplying by M_{jt}, dividing by $Y_{jt} = G(K_{jt}, L_{jt}, M_{jt}) \exp^{\omega_{jt}+\varepsilon_{jt}}$ and taking logs, we get,

$$\log\left(\frac{c_t M_{jt}}{P_t Y_{jt}}\right) = \log\left(\frac{G_M(\cdot) M_{jt}}{G(\cdot)}\right) + \mathcal{E} - \varepsilon_{jt}$$

From this equation, we can recover the elasticity of the production function and the error term. Integrating it, we can recover the production function up to an unknown constant depending on K_{jt} and L_{jt}. Then using the moments generated by the Markov process of productivity, we can recover the remaining constants. Having the production function and the error term, ω_{jt} can be recovered at the firm-level with the difference.

The question for us is, how does the decision of the firms in adopting AI look like? When do they do this? Can we leverage from modeling this decision to identify the effect of adopting AI on productivity.

For example, we can start assuming that firms choose to adopt AI in the current period (whatever the data tell us a period is) before knowing their productivity shock in the current period. In other words, a firm chooses $AI = 1$ iff,

$$\mathbb{E}[\pi | AI = 1] > \mathbb{E}[\pi | AI = 0]$$

A firm would not adopt AI $AI = 0$ if

$$\mathbb{E}[\pi | AI = 1] \leq \mathbb{E}[\pi | AI = 0]$$

Of course, simplifying these inequalities we get that firms adopt AI if

$$\mathbb{E}[\omega_{jt} | AI = 1] > \mathbb{E}[\omega_{jt} | AI = 0]$$

and they do not kwow if

$$\mathbb{E}[\omega_{jt} | AI = 1] \leq \mathbb{E}[\omega_{jt} | AI = 0]$$

It makes sense that firms can flexibly adopt AI because the cost of using generative AI drastically dropped after the release of ChatGPT. Of course, specialized generative AI tools will take longer, but we can restrict the study to the first wave of generative AI, looking at the period close to the release of ChatGPT.

Following our assumptions, we can then have many instruments. We can use k_{jt} as instrument. It will be correlated with the adopting AI, because firms with higher computational power and better IT infrastructure would more easily adopt AI, but because capital is a predetermined input, it is uncorrelated with the current productivity shock.

The same argument works for labor. Skilled labor l_{jt} might be correlated with AI adoption, but it will be uncorrelated with current productivity shock.

Other instruments we could use are the intermediates used last period m_{jt}, the capital and labor used the previous period, productivity used two periods ago.

Data

In terms of data, most individual-level studies rely on experimental data. An experiment could be an avenue for this paper, depending on the question we choose to tackle. The ideal experiment would involve a Canadian enterprise introducing AI assistance tools to its employees. This setup would allow us to control the treatment assignment (who gets access to AI tools) and measure the effects of AI on productivity. To explore the nuances of the effect of AI on productivity, the experiment will require measures of employees' skill levels and tasks with varying levels of difficulty.

Alternatively, an experiment can be conducted at Western University using students as participants. The experiment could involve coding exercises in various programming languages, ranging from popularly used ones like Python and R to less common ones like Fortran or Stata. Popular programming languages should represent a larger share of the gen AI training data set. It is expected that students with access to gen AI would perform better with popular languages such as Python and R, but their performance might decline with less familiar languages like Fortran or Stata. To account for skill levels, proxies such as students' grades in relevant courses could be used. We can also ask their familiarity with each language. The experiment could be conducted at the Ivey Behavioral Lab.

For firm-level data, Statistics Canada (StatCan) conducts several surveys that inquire about firms' use of AI, such as the Canadian Survey of Business Conditions (CSBC) and the Survey of Digital Technology and Internet Use (SDTIU). These surveys can potentially be linked by StatCan to other datasets they maintain, such as the Financial and Taxation Statistics for Enterprises (AFTS). Linking these datasets could help to establish a connection between AI usage and different measures of productivity. This linkage would allow for a more comprehensive analysis of how generative AI impacts firm-level productivity.

Access to the microdata of these surveys and linkages between datasets are restricted to the Research Data Centers (RDCs). There is one in the Social Science building in Western, however, obtaining access to this center and the datasets is a lengthy process.

CSBC

StatCan's Canadian Survey of Business Conditions (CSBC) included a section regarding the use of Artificial Intelligence (AI) by canadian firms in the third quarter of 2024. This survey is conducted quarterly but not all editions of the survey contain these questions.

In the section of this survey, firms are asked if they are planning to use AI and specify which type, including Large Language Models (LLMs). In addition, firms are asked what changes are they planning to implement to make use of AI. Answers include train current staff in AI, Hire staff trained in AI, Use vendor or consulting services to install or integrate AI.

Similar questions were included in the second quarter of 2024. In this edition, firms were asked if they have used AI, whay type, and if the use of AI reduced tasks previously performed by employees.

During the first quarter of 2024, the survey focused on Generative AI. Were business were using, had plans, considered, not considered using AI. Business were also asked if how generative AI created value for them. The options included:

- Accelerate creative content
- Increase automation tasks
- Automate tasks to replace employees
- Improve client or customer experience

among others.

SDTIU

Survey of Digital Techonology and Internet Use is conducted ocassionally. 2023 editions was released in september 2024. The purpose of the SDTIU is to measure the impact of digital technologies on the operations of Canadian enterprises.

The 2023 edition included a section on Artificial Intelligence (AI). The questionnaire specificly asks what type of AI the firms is using, including generative AI. For what purpose the firm uses AI (e.g. marketing, production, administration, logistics, human resources management). In addition, the enterprises are asked the level of usage (e.g. by a few employees, few teams, most or all teams), and the reasons the enterprise has not incorporated AI if they have not.

SIBS

StatCan's Survey of of Innovation and Business Strategy, 2022 (SIBS) include a question regarding the use of artificial intelligence in canadian businesses. According to StatCan, the objective of the SIBS *is to collect information on the strategic decisions, innovation activities, operational tactics and global value chain activities of businesses in Canada.*

The specific question is: *In 2022, did this business use any of the following types of advanced technologies?*

- Artificial Intelligence (AI)

Does not specify which type though (Machine Learning, Large Language Models). However, the last iteration of this survey was collected from January 17 to March 31, 2023. ChatGPT debuted in November, 2022. It was super fresh. Previous iterations of the survey that include the AI question are 2019 and 2017.

Access to this survey would require accessing the Research Data Cente (RDC). There is one in the Social Science Building at Western.

SAT

The Survey of Advanced Techonology of 2022 asks canadian enterprises about their usage of artificial intelligence. However, it does not include generative AI or large language modelds; it focuses on machine learning and natural language recognition among others.

There is not a more recent version available because this is an occasional survey.

References

Brynjolfsson, Erik, Danielle Li, Lindsey R Raymond, Daron Acemoglu, David Autor, Amittai Axelrod, Eleanor Dillon, et al. 2023. "Generative AI at Work," April. https://doi.org/10.3386/W31161.

Czarnitzki, Dirk, Gastón P. Fernández, and Christian Rammer. 2023. "Artificial Intelligence and Firm-Level Productivity." *Journal of Economic Behavior & Organization* 211 (July): 188–205. https://doi.org/10.1016/J.JEBO.2023.05.008.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani. 2023. "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality." *SSRN Electronic Journal*, September. https://doi.org/10.2139/SSRN.4573321.

Gambacorta, Leonardo, Han Qiu, Shuo Shan, and Daniel M Rees. 2024. "Generative AI and Labour Productivity: A Field Experiment on Coding." www.deor.org.cn.

Jiang, Wei, Junyoung Park, Rachel Xiao, and Shen Zhang. 2025. "AI AND THE EXTENDED WORKDAY: PRODUCTIVITY, CONTRACTING EFFICIENCY, AND DISTRIBUTION OF RENTS." http://www.nber.org/papers/w33536.

Kreitmeir, David, and Paul A Raschky. 2024. "The Heterogeneous Productivity Effects of Generative AI the Heterogeneous Productivity Effects of Generative AI *." *SoDa Laboratories Working Paper Series*. http://soda-wps.s3-website-ap-southeast-2.amazonaws.com/RePEc/ajr/sodwps/2024-01.pdf.

Appendix

Lit Review

Czarnitzki, Fernández, and Rammer (2023) Using survey data on German firms, the authors find a positive and significant association between the use of AI and firm productivity. Authors use a production function approach, but they cannot use popular method of production functions because their data is only a cross-section of german firms and a very small panel. Their measure of AI are indicators (e.g. whether they use or not AI) and some measures of intensity of AI usage. The authors define AI broadly because the 2019 survey they use asks firms about usage of AI methods such as natural languange understanding, image recognition, machine learning, and knowledge-based systems, whether these technologies were developed by the firm or a third-party, and the start year of AI usage. They employ IV estimators to account for the possible endogeneity arising from more productive firms are the those investing in AI. The instruments they need are variables correlated with AI use but uncorrelated with productivity (super hard). The instruments they use are sector-level AI use, past innovation expenditure, and resistance to innovation by the firms' workers.

Brynjolfsson et al. (2023) "we study the staggered introduction of a generative AI-based conversational assistant using data from 5,179 customer support agents. Access to the tool increases productivity, as measured by issues resolved per hour, by 14% on average, including a 34% improvement for novice and low-skilled workers but with minimal impact on experienced and highly skilled workers. We provide suggestive evidence that the AI model disseminates the best practices of more able workers and helps newer workers move down the experience curve. In addition, we find that AI assistance improves customer sentiment, increases employee retention, and may lead to worker learning. Our results suggest that access to generative AI can increase productivity, with *large heterogeneity in effects across workers.*"

The tool under study is the Generative Pre-Trained Transformer (GPT) family of large language models (LLM) developed by OpenAI. It monitors customers chats and provides real-time suggestions for how to respond. Agents had the option to ignore the GPT's suggestions.

The effects of productivity are highly heterogeneous across agents. Low-skilled and less-experienced workers improved significantly across different measures of productivity. Access to the AI tool helped newer agents move across the experience curve more quickly. In contrast, there were minimal impacts on the productivity of more-skilled more-experienced agents. Importantly, the authors find evidence that AI assitance may decrease the quality of conversations by the most skilled agents.

AI assistance helped agents communicate more effectively even under hostile conversations, but it also made agenst seem mechanical or unauthentic.

AI assistance helps in the convergence of skill across agents by using data from high-skill workers and helping low-skill workers. Worker contribution to the AI assistance tool is heterogeneous and high-skill workers are not compensated by the data they provide.

Jiang et al. (2025) find that AI exposure is associated with longer work hours and reduced leisure time. The productivity gains in these markets are captured by consumers or firms. Why? Primarily, because AI complements human labor rather than replacing it, according to the authors. Authors use individual-level time diary data from 2004 to 2023 (American Time Use Survey). AI exposure stems froms ChatGPT shock or broader AI evolution.

The author use a simple principal-agent model to explain the mechanism of how AI exposure affects the allocation of individuals' time between work and leisure. They identify three main mechanisms: If AI complements human labor enhancing the marginal productivity of the agent, the increase in marginal producivity would lead agents to work longer hours. AI enhances work monitoring leading to increasing work hours. Lastly, in competitive labor markets where workers have limited bargaining power, workers would not be able to extract the benefits of their increased productivity (higher wages or lower work hours).