Pre-Analysis-Plan: Managers survey

Project: Will Artificial Intelligence Get In The Way of Achieving Gender Equality?

Daniel Carvajal, Catalina Franco, Siri Isaksson

Date: 10.06.2024

Project Summary

This research project aims to investigate the existence of gender differences in the adoption and use of generative AI technologies, specifically ChatGPT. This project seeks four purposes: (i) identify whether a gender gap exists in the adoption of ChatGPT, (ii) explore the underlying mechanisms driving any observed disparities, (iii) evaluate the efficacy of interventions aimed at reducing the gender gap in AI adoption, and (iv) assess the impact of this gap on labor market outcomes. Previous data collections for this project have assessed purposes (i)-(iii), using a sample of university students from NHH Norwegian School of Economics. We find that there is a significant gender gap in use and skills of ChatGPT in students. Moreover, we observe that explicitly allowing the use of ChatGPT in the classroom closes the gap, while on the other hand explicitly forbidding the use of ChatGPT increases the gap.

The current study targets objective (iv) by surveying managers in industries that employ graduates from NHH. First, we use a conjoint-type study to examine whether managers value job candidates that showcase generative AI expertise in hiring decisions, currently and prospectively. Second, we use a hypothetical vignette experiment to assess whether improvements in productivity due to the use of generative AI are rewarded in the workplace. Finally, we collect a series of measures of perceptions and attitudes towards the use of ChatGPT from employers.

Our goal is to document whether the gender disparities in ChatGPT usage that we document with the student sample can contribute to a gender gap in labor market success when students graduate and transition to the labor market.

Participants

The study consists of a survey experiment involving around 1000 individuals employed in companies who hold managerial positions and work in one of the top industries where NHH graduates work: Consulting, Finance, Administration, Accounting. It will be run in June 2024. The survey will be administered through Qualtrics, and to the panel of respondents available to the survey provider company Norstat. All survey answers are anonymous, and the data collected will be used exclusively by the research team.

Survey Design

The survey is structured to take around 7 minutes and includes a conjoint study, a hypothetical vignette, and survey questions on perceptions and attitudes towards the use of generative AI in the workplace. From now on, we will refer to each participant in the study as a manager.

Main outcomes

Value of generative AI in hiring (conjoint study):

Y1. Score given by managers to a hypothetical **current** candidate represented by a profile card. Each manager gives scores to two randomly selected profiles.

Y2. Score given by managers to a hypothetical **prospective** candidate in three years, represented by a profile card. Each manager gives scores to one selected profile.

Value of generative AI in workplace (vignette study)

Y3. Each manager is presented with a hypothetical situation in which two employees worked on a task and their performance on that task defines whether they go into a "promotion track" or not. Workers were allowed to use generative AI and one worker finished the task 25% faster than the other worker. The outcome is a binary variable that takes value 1 if the worker that finished the task faster is selected for the promotion track.

Secondary outcomes

Value of generative AI in hiring (conjoint study):

- Managers are also asked to select between the two current candidates presented to them to be called for an interview.
- Moreover, for the individual selected for the interview, managers are asked how much the participant can negotiate the starting salary if given the position.

ChatGPT Usage and Workplace

- Usage of ChatGPT
- Policies in place in companies
- Influence of ChatGPT in grades
- Value in hiring
- Value in the workplace

Attitudes towards ChatGPT usage

- Advantages
- Disadvantages
- Net benefit

Expectations about the future value ChatGPT (three years)

- Valued in hiring
- Valued in the workplace

Conjoint study

Each manager is presented with two profiles, randomly selected, where the manager must give a score to each candidate, and then select which one will be called for an interview, as well as how much will the candidate be able to negotiate the starting salary. The profiles vary in several dimensions:

- ChatGPT expertise
- Grade and grade distribution
- Gender

The two selected profiles come from a set of 10 hypothetical profiles that represent the following 5 types of workers:

- WHC: Woman with High grades and ChatGPT expertise
- WHN: Woman with High grades and No ChatGPT expertise
- MHC: Man with High grades and ChatGPT expertise
- MHN: Man with High grades and No ChatGPT expertise
- MLC: Man with Low grades and ChatGPT expertise

Vignette experiment

Each manager is presented with a situation in which two employees must work in the same task, and their performance in the task defines whether they are selected for the promotion track at the company. The managers are explicitly told that the use of generative AI is allowed. Both workers have the same output quality, but one worker does the job 25% faster than the other one. Managers are randomly assigned to one out of two main conditions:

- Explicit: managers are given the number of hours that each worker took to finish the task and explicitly told that the faster used ChatGPT and the slower did not.
- Not explicit: managers are given the number of hours that each worker took to finish the task, but it is not explicitly mentioned who used ChatGPT.

The gender of the employees is also randomized, which will be exploited for heterogeneity analysis of the main treatment effect.

Main analyses:

Value of generative AI in hiring (conjoint study)

Current scenario (Y1)

$$y_1 = \beta_{10} + \beta_{11} WHC + \beta_{12} MHN + \beta_{13} MHC + \beta_{14} MLC + X\gamma_1 + \varepsilon$$

Where y_1 is the score given to the candidate, and WHC, MHN, MHC, MLC are dummy variables that take value 1 if the candidate is of each of the respective types, and 0 otherwise. Note that the baseline is type WHN. The comparisons of interest in our analysis correspond to:

1.1. WHC-WHN: given by β_{11}
1.2. MHC-MHN: given *by* $\beta_{13} - \beta_{12}$
1.3. MLC-WHN: *given by* β_{14}

Future scenario (Y2)

$$y_2 = \beta_{20} + \beta_{21} MLC + X\gamma_2 + \varepsilon$$

Where y_2 is the score given to the candidate, and MLC is a dummy variable that takes value 1 if the candidate is of the respective type, and 0 otherwise. Note that the baseline is type WHN. The comparisons of interest in our analysis correspond to:

2.1. MLC-WHN: given by β_{21}

Value of generative AI in workplace (vignette study)

Promotion track to fastest worker (Y3)

$$y_3 = \beta_{30} + \beta_{31} Explicit + X\gamma_3 + \varepsilon$$

$$y_3 = \beta'_{30} + \beta'_{31} Explicit + \beta'_{32} Female + \beta'_{33} Explicit \times Female + X\gamma'_3 + \varepsilon$$

Where y_3 is a dummy variable that takes value 1 if the fastest worker was selected for the promotion track and 0 otherwise, $Explicit$ is a dummy variable that take value 1 if the manager was explicitly informed about who used ChatGPT and 0 otherwise. We intend to first estimate the effect of making the use of ChatGPT explicit on y_3, given by coefficient β_{31} (first equation). We also want to study heterogeneity of this effect when the fastest worker using ChatGPT is a man or a woman, given by the estimated coefficient of the interaction term β'_{33} (second equation).

Hypotheses

Drawing on insights from the growing recent literature on the effects of access to generative AI on productivity across fields (Noy and Zhang, 2023; Brynjolfsson et al., 2023; Peng et al., 2023), as well as recent survey data from employees in Amazon (2024), we expect managers to value the use and signaling of generative AI skills in workers. Thus, we expect positive estimates for comparisons 1.1 and 1.2.

We intend to study whether knowledge of ChatGPT makes a low-grade male candidate comparable to a high-grade female candidate without knowledge of ChatGPT. Therefore, we intend to test whether the estimate for 1.3 and 2.1 is zero.

Finally, we do not specify a direction of the hypothesis in whether making explicit the use of ChatGPT affects positively, negatively or has no impact on the decision to choose the fastest candidate. However, we prespecify our interest in looking for heterogeneous effects according to the gender of the fastest candidate. For example, if the use of ChatGPT is perceived negatively, previous work suggests there might differences in retaliation on the use of ChatGPT by gender of the user (Dehdari et al, 2019).

References

Amazon (2024). Accelerating AI skills: Preparing the workforce for jobs of the future. AWS Reports.

Bimber, B. (2000). Measuring the gender gap on the Internet. *Social science quarterly*, 868-876.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447.

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, *109*(3), 739-773.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (No. w31161). National Bureau of Economic Research.

Dehdari, S. H., Heikensten, E., & Isaksson, S. (2019). What goes around (sometimes) comes around: gender differences in retaliation. Available at SSRN 3378279.

Doshi, A. R., & Hauser, O. (2023). Generative artificial intelligence enhances creativity. *Available at SSRN 4535536.*

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*, 187-192.

OECD (2018). Bridging the digital gender divide: include, upskill, innovate. *Available at: www.oecd.org/digital/bridging-the-digital-gender-divide.pdf*

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.