

Machine discrimination and its solution using social image: Evidence from a trust game*

Kiyotaka Yageta[†]

January 19, 2020

Abstract

As a breakthrough in artificial intelligence (AI), AI has assisted humans in every area of our daily lives. Many economists have discussed the effect of incorporating machine into our work on aggregated indices such as growth, employment, and wage. However, there are only few researches in micro perspective. In this paper, I focused on trust and trustworthiness towards machine which are essential factors smoothing interactions with other individuals. I used inequality aversion utility function developed by Bolton & Ockenfels (2000) to derive the optimal behaviors in a trust game and empirically validate the theories by conducting a laboratory experiment which participants played the game with human and machine. I found in a randomized controlled experiment that although participants decreased both trust and trustworthiness when they played the trust game with machine under no-selection condition, social image reduced the machine discrimination. In the case with selection, on the other hand, those who had low trust and low trustworthiness chose more machine than human as their counterpart of the game, and there is no longer any treatment effect for them. This paper would contribute to the future interactions between humans and machines incorporating a framework of social image.

JEL classification: C72, C91, D63, D91, J71

Keywords: Behavioral economics, Experimental economics, Machine discrimination, Social image, Trust game, Inequality aversion

*The survey in this paper was conducted with financial support from the Japanese Society for the Promotion of Science (JSPS) Grant-in-Aid for Challenging Research (KAITAKU) (grant number: 17H06195). The experiment in this paper was approved by Research Ethics Committee of the university of Tokyo (approved number: 18-285).

[†]Graduate School of Economics, The University of Tokyo. I would like to thank Professor Yoji Inaba for inviting me to Challenging Research (KAITAKU) and funding. For my research, I deeply appreciate Professor Yasuyuki Sawada for supervising my master thesis. I also thank Professor Akihiko Matsui and Professor Michihiro Kandori for giving helpful comments for my theoretical model. In addition, I want to thank Professor Hitoshi Matsushima and Associate Professor Naoko Okuyama for supporting my experiment. Finally, I thank Ren Kurosawa and Masahiro Aono for spending a lot of time helping my experiment. (Email: hijiri.yagekiyo@gmail.com)