

Hajime Shimao

Phone: +1 (765)-409-6391
Email: hajime.fr@gmail.com

website: <https://sites.google.com/site/hajimeshimao/home>
Japanese citizen and U.S. permanent resident

Position

Postdoctoral Fellow, McGill University	2021 - present
Postdoctoral Fellow, Santa Fe Institute	2018 - 2021

Education

Ph.D. Economics, Purdue University	2018
Dissertation: <i>Essays on Structural Econometric Modeling and Machine Learning</i>	
M.S. Economics, Purdue University	2014
M.S. Decision Science, Tokyo Institute of Technology (Concentration: Evolutionary Game Theory, Experimental Economics)	2012
B.A. Psychology, University of Tokyo (Concentration: Decision Theory)	2009

Skills

Machine Learning: Interpretable ML, NLP, Graph Neural Network, Representation Learning, Causal Inference, Predictive Analysis, Scraping, Pattern Mining

Economics Background: Econometrics, Microeconomics, Game Theory, Industrial Organization, Structural Estimation

Programming: Python, R, C/C++, Matlab, Mathematica, Stata

Languages: English & Japanese (fluent); Chinese (beginner)

Research

(Please see the last page for descriptions of selected projects.)

Published Papers:

1. “The Use of AI in Legal Systems: Determining Independent Contractor vs. Employee Status” (with Maxime Cohen et al.) *accepted at Artificial Intelligence and Law*, 2022.
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4013823)
2. “An Observational Study of Health Care Provider Collaboration Networks and Heterogeneous Hospital Cost Efficiency and Quality Outcomes” (with Sebastian Linde) *accepted at Medicine*, 2022.
3. “WikiArtVectors: Style and Color Representations of Artworks for Cultural Analysis via Information Theoretic Measures” (with Bhargav Srinivasa Desikan and Helena Miton) *published at Entropy*, 2022.
(<https://www.mdpi.com/1099-4300/24/9/1175>)
4. “Scale and information-processing thresholds in Holocene social evolution” (with Jaeweon Shin et al.) *published at Nature Communications*, 2020.
(<https://www.nature.com/articles/s41467-020-16035-9>)
5. “Nonconvex Optimization for Fair Regression” (with Junpei Komiyama, et. al) *published at Proceedings of the 35th International Conference on Machine Learning (ICML)*, 2018.
(<http://proceedings.mlr.press/v80/komiyama18a.html>)
6. “Reciprocity and Exclusion in Informal Financial Institutions: An Experimental Study of Rotating Savings and Credit Associations” (with Takehiko Yamato, et. al) *published at PloS One*, 2018.
(<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0202878>)

7. “Strict or Graduated Punishment? Effect of Punishment Strictness on the Evolution of Cooperation in Continuous Public Goods Games” (with Mayuko Nakamaru) *published at PloS One*, 2013.
(<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0059894>)

Working Papers:

1. “Welfare Cost of Fairness and Accountability in Insurance Pricing” (with Fei Huang)
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4225159)
2. “Ensuring Strategic Best-Response Fairness in Machine Learning: Enhanced Color Blind Algorithm” (with Warut Khern-am-nuai and Karthik Kannan) *Revise and Resubmit at Information Systems Research*. (Previously 3rd round Revise and Resubmit at Management Science)
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3389631)
3. “Revisiting the CEO Effect Through a Machine Learning Lens” (with Sung Joo Kim et al.)
(draft coming soon)
4. “Cost of Research and Education Activities in US Colleges - Complementarity, Scalability, and Heterogeneous Efficiency” (with Xiaoxiao Li, Michael Price, and Chris Kempes) *submitted to Science Advances*.
(<https://www.researchsquare.com/article/rs-1318655/v1>)
5. “UnivProd: A university production dataset” (with Michael Price et al.) *under review at Data in Brief*.
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4004591)
6. “Major Complexity Index and College Skill Production” (with Xiaoxiao Li and Sebastian Linde) *Revise and Resubmit at Journal of Human Resources*.
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3791651)
7. “So You Think You are Safe: Implication of Quality Uncertainty in Security Software” (with Warut Khernamnuai and Kirthik Kannan) *Reject and Resubmit at Information Systems Research*. (Previously 3rd round Revise and Resubmit at Management Science)
(https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2621846)
8. “The Past as a Stochastic Process” (with David Wolpert et al.) *Revise and Resubmit at Cliodynamics*.
(<https://arxiv.org/abs/2112.05876>)
9. “Cross-Validation Based Model Selection on Generalized Method of Moments with Application to Dynamic Pricing Model” (with Junpei Komiyama)
(<https://arxiv.org/pdf/1807.06993.pdf>)

Works in Progress:

1. “Modeling Smart Contract Agents with Multi-Agent Reinforcement Learning and Influence Diagram” (with David Wolpert et al.)
2. “Causal Effect of CEO on Compensation and Firm Performance” (with Masahiro Asami)
3. “Dynamical System of Equal Opportunity” (with Vicky Yang et al.)
4. “Optimizing Efficiency of Physician Collaboration Network by Graph Reinforcement Learning” (with Sebastian Linde)
5. “Evolution of Organizational Structure with Information-constrained Agents” (with David Wolpert et al.)
6. “Neural Production Function Approach to Understand Scalability and Complementarity of Multi-objective Organizations” (with Xiaoxiao Li et al.)
7. “Nonparametric Stochastic Process Inference for Empirical Social Science” (with David Wolpert and Jin Hong Kuan)
8. “Private Sector Influence on Federal Reserve Monetary Policy” (with Michael Curran and Craig McMahon)

Teaching Experience

Teaching Assistant (Undergrad Level)

Econ 499 Honors Thesis Course Fall 2015

Teaching Assistant (Master Level)

Econ 510 Game Theory Spring 2016

Mathematical Modeling in Social Science Spring 2009

Teaching Assistant (Ph.D. Level)

Econ 621 Applied Industrial Organization Fall 2016

Econ 631 Industrial Organization Spring 2016

Econ 673 Time Series Econometrics Spring 2016

Other Teaching Experience

Mentor for *Research Experience for Undergraduates (REU)* by NSF 2018-2020

Mentor for *Santa Fe Indian School* 2018-2019

Tutor for three individual students 2007-2011

Tutor for *Scientific Education Group* 2004-2006

Professional Activities

Conference and Seminar Presentations (some serving as Chair and Discussant)

- 2022 Conference on Information Systems and Technology (CIST), Villanova University, VSB Research Day, AAAI/ACM Conference on Artificial Intelligence, Ethics and Society (AIES), ASSA Annual Meeting
- 2021 Southern Economic Association Conference, Purdue University Labor Lunch, Swarthmore College, Western Economic Association International, Villanova University, The Society of Labor Economists, NBER Economics of Education Program Meeting Spring, Midwest Economics Association Conference
- 2020 Southern Economic Association Conference, Western Economic Association International, Eastern Economic Association Conference
- 2019 Conference on Information Systems and Technology (CIST), INFORMS Workshop on Data Science, Annual Workshop on Information Technologies and Systems (WITS), Midwest Econometrics Group Conference
- 2018 Midwest Econometrics Group Conference, North American Summer Meeting of the Econometric Society, 35th International Conference on Machine Learning (ICML), State University of New York at Binghamton, Federal Reserve Board of Governors, Virginia Tech, St. Gallen University, ASSA Annual Meeting
- 2017 Southern Economics Association Conference, Kent State University, Villanova University
- 2016 Purdue Ph.D. Research Symposium
- 2015 Conference on Information Systems and Technology (CIST)
- 2010 Game Theory Workshop, Kyushu University
- 2009 Human Behavior and Evolution Society of Japan, Kyushu University, Research Institute for Mathematical Science, Ryukoku University Kyoto

Workshop Organizer

- 2020 Evolution of Collective Computational Abilities of (Pre) Historic Societies (Santa Fe, NM)
- The Interplay of Large-Scale Impersonal Trends, Big Ideas, and Great Leaders in History (Santa Fe, NM)

Referee

MIS Quarterly, EPJ Data Science, Statistics and Computing

Editorial Service

2010-2011 Editorial Staff of Letters on Evolutionary Behavioral Science

Awards, Grants, and Fellowships

Postdoctoral Research Funding from <i>IVADO</i>	Fall 2021 - Summer 2023
Graduate Scholarship, Purdue University	Summer 2015 - Summer 2017
Fellowship from <i>Japan Student Service Organization</i>	Summer 2012 - Summer 2015

Description of Selected Projects

Welfare Implications of Fairness and Accountability for Insurance Pricing

While the fairness and accountability in machine learning tasks have attracted attention from practitioners, regulators, and academicians, their consequence in terms of stakeholders' welfare is under-explored especially via empirical studies. In this paper, we developed a framework to evaluate the impact of various fairness and accountability regulations on both consumer welfare and firm profit in the context of insurance market. The scope of this paper covers the entire insurance pricing process and associated existing and potential regulations on both cost modeling and pricing, which extends the existing work that mainly focuses on cost modeling (technical prices). We apply our approach to the data of auto insurance market and show that the accountability requirement incurs significant costs on the insurer and consumers. Fairness-aware ML algorithms on cost modeling alone cannot achieve fairness in the market price or welfare, while they significantly harm the insurer's profit and consumer welfare, particularly of females. Our results also demonstrate that the fairness and accountability constraints considered on the cost modeling or pricing alone cannot achieve individual-level fairness (such as the EU gender-neutral insurance pricing) unless we combine the price optimization ban with particular individual fairness notions in the cost prediction.

Neural network based production cost function estimation

In this project that I'm leading, we estimated production cost function of multi-objective organizations, such as universities, hospitals, and cities. While data (such as measurement of costs and related variables) has substantially improved in recent years, most researchers and practitioners still rely on traditional parametric models that lack the flexibility to evaluate the heterogeneity across organizations, and therefore fail to utilize the increasingly available large datasets. To solve this issue, I leverage neural networks with Bayesian optimization to flexibly estimate the heterogeneous cost functions, which in turn generate various useful insights for researchers and practitioners. With help from collaborators specialized in different domains, I have been applying this novel method to important contexts, including higher education industry, healthcare management, and city planning.

Causal Impact of Stakeholder Opinion on FED's FOMC Policy Rate Decisions

Do private stakeholders' preferences about interest rates impact central bank decision making? Anecdotes and some empirical evidence imply that governmental and private sector pressure groups could influence the monetary policy. In this project, coauthors and I examine the relation between the FED's FOMC policy rate decisions and indices that measure the desires of private sector lobbying groups. We first construct a new dataset that collects stakeholders' opinions on the monetary policy by applying modern natural language processing techniques to automatically collected news articles. We then utilize the causal machine learning techniques to estimate the impact of these opinions on the Fed's decision on the interest rate. Controlling for other economic environment, our results suggest that the private sector have significant impact on the monetary policy, while there also exist substantial heterogeneity based on the economic environment. As a side benefit, our dataset will be publicly available so that others can explore various aspect of opinion formations and their impact on policies.

The use of AI in legal systems

In recent years, the use of AI to aid decisions in legal contexts has become increasingly popular. In collaboration with lawyers, I support the investigation of AI usage in the domain of employment law: determining whether a worker is an employee or an independent contractor. My role in this project was i) to explain and help lawyers structure the data in a way that data scientists can work on, and ii) to perform ML analyses. The first task is not easy but extremely important. It requires patient communication with experts in different domains with no quantitative background. Eventually, we structured data for all Canadian court cases related to this legal question between 2002 and 2020, and demonstrated that simple off-the-shelf AI models correctly classify the cases with more than 90% out-of-sample accuracy. The analysis is subsequently deployed at MyOpenCourt.org—a free platform based on our AI models that helps users with legal questions on employment law. This platform has already helped hundreds of Canadians who lost their jobs amid the pandemic. This paper is recently accepted at *Artificial Intelligence and Law*.