

Michael Vinh Pham

linkedin.com/in/michael-vinh-pham
mpham1@swarthmore.edu
mpham8.github.io
www.github.com/mpham8

Education

Sep. 2021 - May 2025 **Swarthmore College**, B.A., Mathematics (3.7 GPA) and Economics (3.8 GPA)
-Sigma Xi Honor Society, 2-year D3 XC/Track Student-Athlete, Swarthmore College Computer Society

Research Assistance

June 2025 - Present **Economics Research Analyst**, *Federal Reserve Bank of New York*
-Dynamic Stochastic General Equilibrium Team in Monetary Policy Division

May 2024 - July 2024
(Full-time) **Economics Research Assistant for Dr. Stephen O'Connell**, *Swarthmore Economics Department*

Sep. 2024 - May 2025
(Part-time) -Created overlapping generations heterogenous agent DSGE model, wrote Python scripts to numerically solve Steady States via Value Function Iteration
-Writing EM Maximization simulation Python script with reduced form model to estimate human capital formation parameters, Simulation Method of Moments with optimization model to estimate parents' perceived human capital formation parameters
-Wrote Python scripts to perform welfare annuity equivalence calculations, pull and analyze IMF GFS data, data visualization
-Read literature and modified Dynare scripts to incorporate human capital investment with dynamic complementarity and self-productivity into macroeconomics models

Jan. 2024 - Present
(Part-time) **Economics Research Assistant for Stephanie Kestelman**, *5th-year Harvard Economics Ph.D. candidate*
-Wrote data preprocessing, machine learning, text analysis Python scripts to conduct sentiment analysis, sequence-to-sequence translation, word embedding unsupervised learning from LA City Planning Commission hearing audio files
-Employed OpenAI API to finetune and run inference to create data from LA housing entitlement data
-Pulled Google Streetview imagery and trained models to perform semantic segmentation Streetview imagery and other computer vision algorithms
-Developed API scripts, web scraping scripts using Selenium/BeautifulSoup to scrape housing data
-Used Stata and Python to merge data, data-preprocessing housing entitlement datasets spanning hundreds of thousands of observations

June 2023 - Sep. 2023
(Full-time) **Machine Learning Research Assistant for Dr. Ryan Kastner and Dr. Curt Schurgers**, *University of California San Diego Computer Science and Engineering Department (REU)*
-Researched, experimented with CV techniques to semantically segment UAV, satellite imagery to track Mangrove Tree deforestation
-Created pipeline to source satellite data, geographic information system processing, feature engineer, run machine learning models with GeoPandas, Pytorch, Hugging Face
-Improved the lab's mangrove satellite identification from 93% to 98% pixel-wise accuracy via experimenting with preprocessing techniques, architecture selection, training methods

Feb. 2023 - June 2023
(Part-time) **Machine Learning Research Assistant for Dr. Xiaodong Qu**, *Swarthmore College Computer Science Department*
-Read 100+ machine learning for EEG classification papers and second-authored systematic review papers accepted to KDD and HCII conferences

Nov. 2022 - May 2023
(Part-time) **Bioinformatics Research Assistant for Dr. Renuka Nayak**, *University of California San Francisco School of Medicine*
-Wrote R, Python Tensorflow scripts to analyze high-dimensional, sparse microbiome arthritis development data
-Developed new machine learning architecture achieving 30, 38, 44 ppt info gain on MCC score over best existing Deep Learning model when predicting Gender, Abx treatment, SKG/WT phenotypes/traits, respectively

Teaching Assistance

Fall 2024, Spring 2025	Social Sciences Quantitative Lab (Stata)
Fall 2024	Intermediate Microeconomics
Spring 2024, Fall 2023	Introduction to Econometrics
Spring 2024	Public Economics
Fall 2022	Introduction to Economics

Other Work Experience

May 2022 - August 2022 (Full-time)	Quantitative Intern , <i>Karner Blue Capital</i>
---------------------------------------	---

Personal

Programming Languages	<i>Expert:</i> Python. <i>Extensive Use:</i> Stata, MatLab, SQL, JavaScript. <i>Familiar:</i> C++/C, R, x86 Assembly
Data/ML Technologies	Pandas, PyTorch, Tensorflow, Hugging Face, SciKit, GIS (QGIS, Geopandas), BeautifulSoup
Other Technologies	Docker, git, AWS, Next.js, Django, React/React Native, SupaBase, HTMX, Tailwind, HTML/CSS
Citizenship	United States
Languages	English (Native), Vietnamese (C1), Spanish (B2), French (A1), Mandarin (A2)

Selected Works

- 2025 **Estimating Intergenerational Effects of Early Age Human Capital Investment** with Stephen O'Connell, *Working Paper*
Should public spending on early-age human capital formation be protected during a fiscal consolidation? The question is relevant to countries throughout the global income distribution. We address it using a DSGE with overlapping generations in which parental inputs are combined with public resources to produce skills that become productive when children reach adulthood. Private inputs are subject to a borrowing constraint, and potentially also to parental misperceptions of the human capital production function. We show how to calibrate a DSGE like ours to match these distortions in the data. We then use stochastic simulations to quantify the impact of reductions of public support to early-age human capital formation on lifetime income and wage inequality within the affected cohorts, and on the intergenerational transmission of inequality.
- 2025 **Barking up the Wrong Tree? Multifamily Housing and Tree Loss: Evidence from Los Angeles**, *Working Paper*
This paper evaluates how the construction of multi-family properties in Los Angeles affect the neighborhood characteristic nearby as defined as greenery coverage. I create address-level panel data by employing computer vision techniques applied to street-level imagery and use difference-in-differences estimators. I find that multi-family development has no negative effect on nearby tree and greenery coverage and may even contribute to a modest increase, with estimates indicating a statistically significant 2.3% rise in total greenery (trees and shrubs) coverage.
- 2024 **How Universal Healthcare Affects Individuals' Choices: The Case of Seguro Popular**, *Advanced Econometrics Final Paper*
This paper assesses the effects of Seguro Popular, a universal healthcare program implemented in Mexico, on individuals' labor force participation, family planning, and the take-up of womens' preventative cancer screenings. I exploit variation in roll-out of municipalities to find the intent-to-treatment effect using event study estimators with differential timing and use a machine learning based method, Matrix Completion, to assess the robustness of these results. I find that the implementation of Seguro Popular increases the target demographic for Mammogram take-up of such exams, has no significant effect on take-up of Papanicolaou Tests, and find some evidence to support the claim that Seguro Popular increased the amount of children women wanted to have. Finally, Seguro Popular had a negative effect on labor force participation, most notably among working males.
- 2023 **A Review of Machine Learning Algorithms for EEG Datasets** with Nathan Koome Murungi, Xufeng Caesar Dai, and Xiaodong Qu
Accepted: 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (August 2023).
In this paper, we present a systematic literature review that explores the utilization of machine learning (ML) algorithms for analyzing datasets from Electroencephalography (EEG) based Brain-Computer Interfaces (BCIs). Our primary aim is to provide computer science students with a comprehensive and accessible overview of the role of machine learning in EEG analysis. By synthesizing and organizing recent research from 2020 onwards, our objective is to empower the target audience to develop a solid foundational understanding of the current state of ML-EEG research. Through this work, we intend to enhance the accessibility and comprehension of ML-EEG studies and contribute to advancing BCI technology.