

Statement of Purpose

It is my desire to pursue a Ph.D. in Economics as part of my long-term professional goal of becoming a professor. I have a strong interest in applied economics and econometrics; I have aligned my academic pursuits as such.

After graduating from college, I spent three years working for MSH China¹ to meet my student loan obligations. Having achieved this, I began the next phase of my education – a Master degree in Actuarial Studies at Monash University, Australia. My plan was to get a degree to facilitate my future career, but it was during this time that I discovered my passion and talent for academic research. The first trigger happened in the Introductory Econometrics course taught by Professor Farshid Vahid². When I tried to derive the formulas of the estimated coefficient of OLS and checked my solutions with the textbook (Wooldridge 2016). Our equations differed. I was confident about my answers. Therefore, I brought this to Farshid and even he was surprised by these typos. He wrote an email to Wooldridge to inform him, this resulted in a thank you from Wooldridge, and some bonus points from Farshid. With this experience, my confidence and interests both increased exponentially. The more I studied Econometrics, the more I started to enjoy it. I decided to transfer to Master of Applied Economics and Econometrics after the first semester.

In the year that followed, I achieved the best performance in seven out of eight courses. In April 2018, I was chosen as one of the four representatives to participate the Econometrics Game held in Amsterdam University. In two days' time, we wrote a paper to find “the multiplier between the detrimental effect of unemployment on the wellbeing of the individual who is unemployed and the total effect of that person's unemployment on the wellbeing of the group as a whole” (P.Frijters 2018). I can still feel the passion every time I remember how we worked together under great pressure discussing ways of measuring “wellbeing”, constructing the ordered probit model, etc. I was impressed by the breadth and depth of the top three papers given the tight schedule, while also feeling deeply proud for ourselves. Our skills and knowledge open the door to a unique perspective of this world. Our work may affect from an individual's happiness to a country's development. Communicating with the PhD students from other top universities, I realized that the more important thing for doing research is

¹MSH China is the Asia pacific headquarter of MSH INTERNATIONAL. Founded in 1974, MSH INTERNATIONAL is a world leader in the design and management of international healthcare solutions. <http://www.mshchina.com/>

²Farshid Vahid is a professor in the Department of Econometrics and Business Statistics at Monash University.

the ability of applying the knowledges we learned. And the only way of achieving this is by practice.

Coming back from the game, I delved into my own research project which was supervised by Professor Di Cook³. The central question was: can we train the computer to read residual plots? Residuals plots contain more information than single statistics (Anscombe 1973), thus are a primary means to diagnose statistical models. Majumder, Hofmann, and Cook (2013), has proven hypothesis testing of human observations of visualized plots to be a valid methodology. In the hope that previously manual process can be aided and supplemented by computers, we built a binary classifier using a deep learning algorithm which detects linear relationships from no structure in the scatter plots. Having trained a satisfactory model using simulated data, we tested it against the same dataset used by Majumder et al. The human outperformed both t-test and the classifier. However, an interesting aside discovery was that our deep learning model always behaved similarly to t-test. Theoretically, t-test is the uniformly most powerful test under such experimental settings (J.Neyman and E.S.Pearson 1933). The fact that the deep learning model approaches t-test gives hope that it may be able to find the unknown best test in more complicated situations. We then extended our study to test heteroscedasticity from homoscedasticity using the binary classifier. To provide reference, another small human experiment was conducted via online questionnaire and a specific form of white test (White 1980) was employed. In this experiment, the classifier achieved higher accuracy than the white test and humans. This makes hypothesis testing of computer vision of data plots more interesting. Now we are working towards the publishment of these results.

There were tough moments during the research when I needed to learn many new methodologies and theories on a tight deadline. However, the sense of achievement of overcoming obstacles and the enjoyment of solving puzzles confirmed my decision to pursue my career in academia.

After completed the master degree, I accepted the offer from Professor Heather Anderson⁴ to work on a challenging and interesting project as a research assistant. In this project, we investigated the properties of Lasso⁵ estimates(Tibshirani 1996) in the high-dimensional predictive regression in the presence of cointegration(Engle and Granger 1987). More specifically, we studied the consistency of Lasso for estimating the cointegrating vectors and

³Dianne Helen Cook is an Australian statistician, the editor of the Journal of Computational and Graphical Statistics, and an expert on the visualization of high-dimensional data.

⁴Heather Anderson is the Maureen Brunt Professor of Economics and Econometrics, in the Department of Econometrics and Business Statistics at Monash University.

⁵Least Absolute Shrinkage and Selection Operator

applied this methodology to stock return forecasting and bond excess return forecasting.

Meanwhile, I am assisting the learning and teaching team in Monash University to analyse the treatment effect of the education innovations including the live streaming used in the lecture etc.

References

Anscombe, FJ. 1973. “Graphs in Statistical Analysis.” *The American Statistician* 27 (1): 17–21.

Engle, Robert F, and Clive WJ Granger. 1987. “Co-Integration and Error Correction: Representation, Estimation, and Testing.” *Econometrica: Journal of the Econometric Society*. JSTOR, 251–76.

J.Neyman, and E.S.Pearson. 1933. “IX. on the Problem of the Most Efficient Tests of Statistical Hypotheses.” *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 231 (694-706). The Royal Society: 289–337. doi:10.1098/rsta.1933.0009.

Majumder, Mahbulul, Heike Hofmann, and Dianne Cook. 2013. “Validation of Visual Statistical Inference, Applied to Linear Models.” *Journal of the American Statistical Association* 108 (503). Taylor & Francis Group: 942–56.

P.Frijters. 2018. “The Central Question of the Econometric Game.” <http://econometricgame.nl/econometric-game-2018/>.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 267–88.

White, Halbert. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica: Journal of the Econometric Society*. JSTOR, 817–38.

Wooldridge, Jeffrey M. 2016. *Introductory Econometrics: A Modern Approach, 6th Edition*. Cengage.