Statement of Purpose

It is my desire to pursue a PhD in Economics at University California, San Diego as part of my long-term professional goal of becoming a professor. I have a strong interest in applied econometrics.

After graduating from college, I spent three years working for MSH China[1] to meet my student loan obligations before beginning a Master's degree in Actuarial Studies at Monash University[2]. My goal was to facilitate a future career as an actuary. However, I developed a keen interest in econometrics and academic research. This passion stemmed from three mistakes I discovered in the equations in the textbook (Wooldridge 2016) while doing extracurricular reading for the Introductory Econometrics course taught by Professor Farshid Vahid. Surprised by these typographical errors, Professor Vahid wrote to Wooldridge, which resulted in a thank you from him, public praise, and extra course credit points. Encouraged by this experience, I decided to pursue my interest and transferred to a Master's degree in Applied Economics and Econometrics.

In the year that followed, I received the Monash Business School Student Excellence Award for achieving the highest mark in seven out of eight courses. In April 2018, I participated the Econometric Game[3] as one of the four representatives of Monash Business School competing with the best PhD and master students from other prestigious universities. The research question for competition was around the detrimental effects of an individual's unemployment on that individual's happiness as well as on a group's wellbeing. After reviewing the relevant literature for the empirical support on selecting explanatory variables, we constructed an ordered probit model with the raw responses to the survey item that elicits respondents' overall life satisfaction as the dependent variable. With the assumption of the homogeneous spillover effects amongst individuals in each group, we estimated the multiplier between the the effects on an individual and a group. The potential simultaneity bias was discussed. This great experience of conducting rigorous research provided me important new insights into how unemployment affects well-being, and greatly stimulated my interest in empirical research in economics.

This project was supervised by Professor Dianne Cook. The central research question was:

[1]MSH China is the Asia-Pacific Headquarter of MSH INTERNATIONA, a world leader in the design and management of international healthcare solutions. See http://www.mshchina.com/

[2]Monash University is a public research university based in Melbourne, Australia.

[3]The Econometric Game is hosted by the University of Amsterdam. See http://econometricgame.nl/

can we train the computer to read residual plots? Residuals plots contain more information than single statistics (Anscombe 1973). It has been shown that hypothesis test based on human observations of visualized plots is valid (Majumder, Hofmann, and Cook 2013). In the hope that human evaluation can be aided and supplemented by computers, I built a binary classifier using a deep learning algorithm to detect linear relationships from no structure in a scatter plot. Having trained a satisfactory model using simulated data, I tested it against a large unseen test set, as well as the same dataset used by Majumder et al. An interesting discovery was that the power $(1 - \beta)$ of the deep learning model was always close to the conventional t-test holding the Type I error (α) constant. Given that the t-test is the known uniformly most powerful test under such experimental settings (Neyman and Pearson 1933), this finding gives hope that the deep learning model may be able to approach the unknown best test in more complicated situations. The study was then extended to test the null of homoscedasticity against heteroscedasticity using the binary classifier. A small experiment of human observations was conducted via online questionnaire and a specific form of the White test (White 1980) was employed to provide a reference level of the test accuracy. In this experiment, the classifier achieved much higher accuracy than both the White test and humans. Our research extended the use of computer vision on hypothesis testing and provided evidence of its validity when the structure in the plots is very specific. A draft paper is currently being written to be submitted to the journal: Statistical Analysis and Data Mining.

After completing the master degree, I was excited to accept an offer from Professor Heather Anderson as a research assistant to work on a paper that studying the high-dimensional predictive regression with the Lasso estimator (Koo et al. 2016). In this project, I reviewed literature to investigate the compatibility (or restricted eigenvalue) condition and its implication of choosing the tuning parameter λ for the ℓ_1 norm to achieve the prediction optimality, while taking into account the potential consequences in consistency for variable selection of the Lasso (Bühlmann and Van De Geer 2011). Comparison between the Lasso, the adaptive Lasso and the group Lasso is under considering. In addition, the out-of-sample mean squared error of forecasting GDP growth and inflation using the Lasso on 146 economics variables is tested against other approaches including an autoregressive model and a principal component analysis. The potential co-integrating relationships in the selected variables are being studied. I find that it is intriguing to observe the series chosen by this data-driven methodology and seek possible economic theory-based explanations.

Meanwhile, I am co-authoring a paper with the learning and teaching team at Monash University that measuring student levels of perceptions of live-streaming, a new technology

implemented in the lectures. Our study adapted the CRiSP questionnaire (Richardson et al. 2015), positioned as part of an institutional integrative framework. A combination of factor analyses and survey item response theory was employed to examine item and scalar equivalence in order to validate the instrument. Results revealed three scales that reliably measure student perceptions of live-streaming technology: Acceptance of Live-Streaming Technology, Usability of Live-Streaming Technology, and Confidence in Live-Streaming Technology. The adapted and validated instrument demonstrates an improvement of psychometric measures when compared to the CRiSP questionnaire. The scale's validation provides a reliable instrument to understand the how higher-education students perceive live-streaming technology. Our paper presents the overarching framework and describes the adapted and validated CRiSP instrument. Implications for future use of the instrument will also be discussed.

References

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

Bühlmann, Peter, and Sara Van De Geer. 2011. *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Science & Business Media.

Koo, Bonsoo, Heather M Anderson, Myung Hwan Seo, and Wenying Yao. 2016. "High-Dimensional Predictive Regression in the Presence of Cointegration."

Majumder, Mahbubul, 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.

Neyman, J., 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.

Richardson, Alice M, Peter K Dunn, Christine McDonald, and Florin Oprescu. 2015. "CRiSP: An Instrument for Assessing Student Perceptions of Classroom Response Systems." *Journal of Science Education and Technology* 24 (4). Springer: 432–47.

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.