

# BIAN, YONG

909 University Ave., 118 Professional Bldg., Columbia, Missouri 65211  
+1(573)-200-4423  $\diamond$  ybzkf@mail.missouri.edu

## EDUCATION

---

**University of Missouri, Columbia** *May 2021(expected)*

Ph.D. in Economics GPA:3.72

Dissertation: Essays on Machine Learning Applications in Economics: Causal Inference and Prediction

**The Chinese University of Hong Kong** *Dec 2014*

M.S. in Economics

Honors/Awards: Excellent Performance Award (Ranked No. 1 in Applied Econometrics class); Dean's List

**Central University of Finance and Economics, Beijing, China** *Jun 2013*

B.S. in Statistics

Honors/Awards: National Research/Survey Analyst (National Certificate); Summer Volunteer (2011) in Yu Huan, Zhejiang Province

## RESEARCH PAPERS

---

### **A Model Discussion and DML Application–Based on California Math Curriculum Causal Effect Analysis**

- When applying machine learning techniques with high dimensional data in causal effect estimations, the regularization bias causes its estimator not consistent in square root rate. Double/Debiased Machine Learning (DML) successfully obtains a properly converged estimator. In this paper, I clearly explain what is DML to a broader audience group and use it to a causal effect application. Moreover, I take a bootstrap procedure to improve the DML standard error reporting system. As an extension to the existing studies on how curriculum materials affect student achievement, my work focuses on the effect of California Math, an elementary math textbook, on student achievement and compare the model results of DML, Kernel Matching and Ordinary Least Square (OLS). In my study, the DML estimators avoid the possible misspecification bias from linear models and obtain statistical significant estimators which beat the Kernel Matching results.

### **Causal Inferences on Young Economics Professors' Salaries (Job Market Paper, submitted)**

- In this paper, we analyze the causal effects of three interesting factors on young economics professors' salaries: gender, PhD graduation school rank, and undergraduate major. The dataset used is novel, containing detailed research productivity measures and other demographic information of young economics professors from the top 50 public research universities in the United States. We apply double/debiased machine learning (DML) to the models and obtain consistent estimators under high-dimensional control variable space. By tracking the first 10 years of their professional work experience, we find that there barely exists causal effects on salaries from the above three factors in most of the years when controlling for research productivity and background, neither economically nor statistically. However, the gender effect on experience in year 7 is both statistically and economically significant. For both PhD graduation school rank and undergraduate major, the estimates for experience years 7–9 are large in magnitude, although there is a lack of statistical significance in almost all the experience years. Next, based on what have been found, we discuss possible economic mechanisms and reasons.

### **Paper Quality and Gender Bias Based on Text Analysis**

- In this paper, we build machine learning and simple linear models in predicting paper outcomes and discuss the gender bias associated with it. A novel set which contains paper text content and its associated H-index, gender and other information, collecting from recent years published journals is used. We apply tf-idf vectorization and other NLP tools to transfer text content into numeric features as model inputs; and take H-index as model outcome variable. We give adequate descriptive analysis on paper H-index by gender and discuss the possibilities of gender causal effect. We find when using paper text features to predict H-index, the prediction power is about 60% of classification and about 44 root mean square error of regression. Moreover, when controlling for paper text the gender causal effect hardly exist. As long as

the paper contains similar text, gender would not cause its H-index change. Besides, we give real world meanings associate with the models.

## PAPERS IN PROGRESS AND OTHER RESEARCH

---

### **Health Impacts of Household Sanitation Conditions in Rural China (Working in Progress)**

- Investigate the household sanitation conditions impact on elder people's health.
- Focusing on living standards of people from rural China.
- Causal effect analysis under high-dimensional controls.
- Discuss possible polices that could help improving health conditions, living standards and longevity in rural areas.

### **Recurrent Neural Network for High Dimensional Probability Estimation in Economic Modelings (Working in Progress)**

- Use recurrent neural network to estimate probabilities of autocorrelated series; build 3-D input and 3-D output prediction model.

### **Forecasting Stock Returns Using Machine Learning Methods**

- Built Lasso, Elastic net and Ridge models in predicting stock returns.

### **An Exploration of Logit Model - Combined with Ridge based on drug and alcohol data**

- Built high dimensional binary prediction model

## WORK EXPERIENCES

---

### **Department of Economics, University of Missouri**

*Research Assistant*

*Jun 2017 - Aug 2017*

- Project: "Modeling and Extrapolating Wheat Producer Support Using Income and Other Factors."
- Coded in R for data cleaning ("dplyr"); Penal data analysis, built fixed effect models; made graphs ("ggplot")

*Teaching Assistant*

*Aug 2015 - Present*

- Held tutorials, exam reviews and office hours;
- Three semesters of **Head TA**: as a class organizer, helped the professor with administrative tasks, bridge between professor and students.
- TA courses: Econ1015: Principle of Macroeconomics; Econ1014: Principle of Microeconomics; Econ 3229: Money, Banking and Financial Markets; Econ 4353: Intermediate Macroeconomics; Econ 4371/7371: Introductory Econometrics; Econ 4370/8370: Quantitative Economics.

### **Beijing Haidian Dist. Ministry of Statistics, Beijing, China**

*Summer Intern, Data Analyst*

*Jun 2013 - Aug 2013*

- Cleaned survey data, built linear regression models, statistic tables and helped completed 50 pages salary report; assisted managers in other administrative tasks.

## CONFERENCE

---

**Missouri Valley Economic Association 56th Annual Conference**

*Oct 2019*

A Double/Debiased Machine Learning Application in California Math Curriculum Treatment Effect Analysis

## SOFTWARE AND OTHER INTERESTS

---

**R, Python, SQL** Kaggle competitions; certified Coursera classes(deep learning) and Udemy classes(NLP)

## REFERENCES

---

David M. Kaplan (Advisor)

Associate Professor, Department of Economics, University of Missouri

kaplandm@missouri.edu 573-884-8005

Peter Mueser (Committee member)  
Chancellor's Professor, Department of Economics, University of Missouri  
mueserp@missouri.edu 573-882-6427

Saku Aura (Committee member)  
Associate Professor, Department of Economics, University of Missouri  
auras@missouri.edu 573-882-6073

X. H. Wang (Teaching Assistant Supervisor)  
Professor, Department of Economics, University of Missouri  
wangx@missouri.edu 573-882-4954