XIAO WU

170 Brookline Ave, Unit 919, Boston, MA 02215
(+1) 617-513-2976 ⋄ wuxiao@g.harvard.edu

EDUCATION

Harvard University *Cambridge, MA*
Ph.D., Biostatistics *September 2017 - Present*
Dissertation: Causal Inference for Spatial-temporal Data
Committee: Dr. Francesca Dominici, Dr. Jose R. Zubizarreta, Dr. Danielle Braun

Harvard T.H. Chan School of Public Health *Boston, MA*
M.S., Biostatistics *September 2015 - May 2017*

Bachelor of Science in Mathematics *September 2011 - July 2015*
Peking University, Beijing, CHN

Bachelor of Laws *September 2011 - July 2015*
Peking University, Beijing, CHN

ACADEMIC AWARDS & HONORS

American Statistical Association Student Paper Award *2019*
Statistics and the Environment Section

American Statistical Association Student Travel Award *2019*
Biopharmaceutical Section Regulatory-Industry Statistics Workshop

Summer Institute in Statistics for Big Data Scholarship *2017*
University of Washington

1st Prize of the National Mathematics Contest *2009*
The Chinese Mathematical Society (CMS)

RESEARCH EXPERIENCE

Harvard T.H. Chan School of Public Health *Boston, MA*
Statistical Researcher; Mentor: Dr. Francesca Dominici *June 2017 - Present*

- Developed causal inference methods to estimate the exposure response functions for non-binary exposures for very large administrative databases
- Developed measurement error adjustment methods based on regression calibration for the causal estimates in air pollution statistical models
- Developed scalable computing tools to analyze massive spatial-temporal data

Harvard Business School *Boston, MA*
Research Associate; Mentor: Dr. Lauren Cohen *July 2016 - March 2017*

- Conducted quantitative research in life settlement market to provide a gold standard investment guideline for stakeholders
- Determined the objective internal rate of returns for life settlement investors in various settings using sampling and prediction methods

Stanford University School of Medicine *Stanford, CA*
Statistical Researcher; Mentor: Dr. Ying Lu *June 2014 - August 2014*

- Developed a novel efficient cross-sectional design with a short-term follow-up to rapidly establish the prognostic validity of new biomarkers
- Built an innovative Markov model to represent the disease incidents for study design and evaluated the finite-sample and asymptotic properties of the model

INDUSTRY EXPERIENCE

Google LLC *Sunnyvale, CA*
Data Scientist Intern; Mentors: Drs. Li Pan, Meeyoung Park *May 2019 - September 2019*

- Developed a machine learning model-based framework to design and analyze network interference experimentation
- Conducted data mining for large size Google Cloud user collaboration networks to better understand the network structures
- Promoted causal thinking in analytic practice to answer meaningful product/business questions with scientific rigor

Sanofi Genzyme *Cambridge, MA*
Biostatistician Intern; Mentor: Dr. Yi Xu *June 2017 - August 2017, February 2019 - May 2019*

- Developed a novel Bayesian clinical design with the goal to optimize interim analysis timing for Bayesian adaptive commensurate designs
- Proposed Bayesian methods to utilize historical adult trial data in pediatric study designs in rare disease research
- Conducted propensity score matching to pair subjects in multiple clinical trials and analyzed matched survival data

McKinsey & Company *Beijing, China*
Part-time Analyst; Mentor: Dr. Jie Cheng *April 2015 - July 2015*

- Participated in designing dynamic pricing and promotion strategies for a leading Chinese e-commerce company
- Compared efficiency of various pricing forecasting models

Peking University Clinical Research Institute *Beijing, China*
Data Analyst; Mentor: Prof. Chen Yao *February 2014 - June 2014*

- Participated in clinical data analysis to address pathological problems
- Conducted statistical modeling and testing using SAS programming
- Drafted statistical reports for commercial clients

TEACHING EXPERIENCE

Harvard T.H. Chan School of Public Health *Boston, MA*
Teaching Fellow, Theory and Methods for Causality II; Instructor: Dr. Andrea Rotnitsky *Fall 2019*
Teaching Fellow, Introduction to Statistical Genetics; Instructor: Dr. Martin Aryee *Fall 2019*
Teaching Fellow, Applied Bayesian Analysis; Instructor: Dr. Lorenzo Trippa *Fall 2018*
Teaching Fellow, Applied Survival Analysis; Instructor: Dr. Rui Wang *Spring 2017*

Harvard T.H. Chan School of Public Health *Boston, MA*
Guest Lecturer, Computing for Big Data - Working with Medicare Data *December 2018*

Harvard Medical School *Boston, MA*
Guest Lecturer, An Introduction to Propensity Score Methods *September 2018*

PUBLICATIONS

Journal Articles

1. **Wu, X.**, Braun, D., Kioumourtzoglou, M.A., Choirat, C., Di, Q. and Dominici, F., 2019. Causal inference in the context of an error prone exposure: air pollution and mortality. The Annals of Applied Statistics, 13(1), pp.520-547.

2. **Wu, X.**, Xu, Y. and Carlin, B.P., 2019. Optimizing Interim Analysis Timing for Bayesian Adaptive Commensurate Designs. Statistics in Medicine, (accepted), pp.1-24.
   *** Winner of 2019 American Statistical Association Student Poster Competition**

3. Won, J.H., **Wu, X.**, Lee, S.H. and Lu, Y., 2017. Cross-sectional design with a short-term follow-up for prognostic imaging biomarkers. Computational Statistics & Data Analysis, 113, pp.154-176.

Submitted Manuscripts

1. **Wu, X.**, Mealli, F., Kioumourtzoglou, M.A., Dominici, F. and Braun, D., 2018. Matching on Generalized Propensity Scores with Continuous Exposures. arXiv preprint arXiv:1812.06575.
   *** Winner of 2019 American Statistical Association Student Paper Competition**

CONFERENCE PRESENTATIONS

1. Causal effects of long-term $PM_{2.5}$ exposure on all cause mortality, Harvard Data Science Initiative 2019 Conference, Boston, MA.

2. Optimizing Interim Analysis Timing for Bayesian Adaptive Commensurate Designs, ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop (BIOP) 2019, Washington, D.C.

3. Matching on generalized propensity scores with continuous treatments, Joint Statistical Meeting (JSM) 2019, Denver, CO.

4. Matching on generalized propensity scores with continuous treatments, Atlantic Causal Inference Conference (ACIC) 2019, Montreal, QC, Canada (**Invited**).

5. Causal Inference Challenges in Air Pollution Research, Atlantic Causal Inference Conference (ACIC) 2019, Montreal, QC, Canada (**Discussant**).

6. Statistical methods for pooling categorical biomarkers from multiple studies, Joint Statistical Meeting (JSM) 2018, Vancouver, BC, Canada.

7. Causal inference in air pollution epidemiology using generalized propensity score matching, Harvard/MIT ACE Center Science Advisory Committee (SAC) Meeting 2018, Boston, MA (**Invited**).

8. Matching on generalized propensity scores with continuous treatments, European Causal Inference Meeting (EuroCIM) 2018, Florence, Italy.

9. Causal inference in the context of an error prone exposure: air pollution and mortality, International Chinese Statistical Association (ICSA), Applied Statistics Symposium, 2018, New Brunswick, NJ (**Invited**).

10. Causal inference in the context of an error prone exposure: air pollution and mortality, Eastern North American Region (ENAR) International Biometric Society Meeting 2018, Atlanta, GA.

11. Methods to estimate causal effects adjusting for confounding when an ordinal exposure is mismeasured in the context of air pollution, Harvard/MIT ACE Center Science Advisory Committee (SAC) Meeting 2017, Boston, MA (**Invited**).

PROFESSIONAL ACTIVITIES

Journal Peer Reviewer
Biometrics, Biometrical Journal, Statistical Sinica, Health Services and Outcomes Research Methodology, Atmospheric Environment, Harvard Public Health Review

Invited Speaker
Harvard Public Health Symposium for Young Leaders in China

Session Chair
Recent Advances in Nonparametric Statistical Methods, Joint Statistical Meeting (JSM) 2018

Biostatistics Consultant
Biostatistics Student Consulting Center, Harvard T.H. Chan School of Public Health

Legal Consultant
Legal Aid Association, Peking University Law School

TECHNICAL SKILLS

Programming Languages	R, Python, SAS, SQL
Software & Tools	Tensorflow, Stan, R Studio, Matlab, Github, Latex
Certificates	SAS Base and Advanced Programming