

Wu Zhu

3737 Chestnut Street, Philadelphia, PA, 19104
Github Website: <https://zhuwu2012.github.io/>

+1 (215)-285-8924
zhuwu@sas.upenn.edu

RESEARCH INTERESTS	Asset Pricing, Machine Learning, Big Data, Macroeconomics, Empirical Corporate Finance, and the Chinese Economy	
EDUCATION	University of Pennsylvania	2021
	<i>Ph.D in Economics</i>	
	<i>Thesis Title: Networks in Finance, Macroeconomics, and Machine Learning</i>	
	Wharton School, University of Pennsylvania	2021
	<i>M.A in Statistics</i>	
	Machine Learning, Portfolio Management, and Networks.	
	CCER, Peking University	2016
	<i>M.A Economics</i>	
	University of Science and Technology, Beijing	2009
	<i>Bachelor in Material Physics</i>	
HONORS AND AWARDS	<ul style="list-style-type: none">• CFRC Best Paper Award, China Financial Research Conference, 2021• XiYue Best Paper Award, Chinese International Conference in Finance, 2021• Finalist of the Best Ph.D. Paper, Mid-West Financial Association, 2020 (MFA, US)• Finalist of the Best Paper in Investment, Annual Financial Management Association, 2020 (FMA, US)• Wharton Mack Institute for Innovation Research (Machine Learning, Networks, and Asset Pricing, 2020)• Wharton Global Initiatives Research Grant (2018, 2019)• Meritorious Winner, Mathematical Contest Modeling, United States (2008)• First Prize, Chinese National College Mathematical Competition of Modeling (2007)• First Prize, the College Mathematical Olympic of China (2006, 2007)	
RELEVANT INTERNSHIP	<ul style="list-style-type: none">• Jun 2018- Sep 2018 Ph.D Summer Intern in International Monetary Fund (IMF) Project: Machine Learning, Investor Sentiment, and Credit Market Crashes To understand how investor sentiment affects credit and stock market. Collect data on real credit structure (yield curves) across countries as early as 1910 (more than 40 countries). Use professional forecasts, newspapers, and media coverage to estimate the forecasted credit structure. Construct the forecast error or disagreement to reflect the investment sentiment. Use the ARIMA model and find the forecast error can well predict the credit and stock market return next quarter (with $R^2 \approx 0.3$).• Jan2014 - Jun2014 R.A. Counsellor Office of the State Council, Central Government, China.	
SKILLS	Python, R, Stata, SQL, Google BigQuery, Clustering Computation	
	Big Data: Structured or Unstructured Data.	
	<ul style="list-style-type: none">• China Business Registration Data (used in papers 2,3, 6): it covers all firms registered in China (text format, 80 million firms at 2019, more than 200GB, long history of 1950-2020, 2 years of data cleaning), and includes firms' basic	

information, shareholders, investments, managers, and historical updates.
Trace back all historical information and construct dynamic firm-to-firm equity-holding network in China.

Matched with various datasets like Annual Survey, National Survey etc.

- Google Patent Dataset (used in papers 4,5,7): it covers all patents issued between 1911-2018 (html format, 14 million patents, > **50GB**) and includes patent application, issuance, transaction, and citation.
Use patent transaction to trace original owner of each patent.
Use Who owns whom to trace the subsidiary-parent relationship.
Develop an algorithm to match firms in patent datasets with firms in CRSP.
Boosting matching rate from 27% to 48% compared to previous work.
- Financial Datasets: Who Owns Whom, I/E/B/S, Thomas Reuters Institutional investors, Raven Packs News Analytics, CRSP (used in papers 5,8,9, >**1TB**)

Causal Inference: see identification of IV, match, DID in papers 2, 3, 5, 6.

Time series: high dimensional dynamic state space model, bayesian Statistics, kalman filter, PCA, dynamic factor models (used in papers 5,7)

Machine Learning: shrinkage, quantile learning, boosting, and reinforcement learning (used in papers 8,9)

PAPERS UNDER REVIEW

1. with R. Vohra, Y. Xing, “The Network Effects of Agency Conflicts” (under review, [preprint](#))
 - *Talk: Winter Meeting of Econometric Society, Dec2020, NSF 6th Annual Conference in Networks, U Chicago, Mar2020.*
 - *A general framework with agents’ strategic interactions in an equity-holding network, systematically examine the role of various frictions within firms in amplifying and propagating the shock, explore the macro implications of agency conflicts.*
2. with Y. Shi, R. Townsend, “Tiered Intermediation in Business Groups” (under review, [preprint](#))
 - **The Finalist of the Best Ph.D Paper, MFA 2020**
 - *Talk: CICF 2021, AEA 2020, MFA 2020, IMF 2019, NAMES 2019, MIT 2019 etc.*
 - *An empirical paper with big data, construct a Bartik IV to identify that equity-holding network plays a role of tiered financial intermediation, identify the significant indirect effect of the bank lending through ownership network.*
3. with F. Allen, J. Cai, X. Gu, J. Qian, L. Zhao, “Ownership Networks and Firms Growth - What do Forty Million Firms Tell us about Chinese Economy?” ([preprint](#))
 - **Best Paper Award, China Financial Research Conference 2021**
 - *Talk: CICF2021, AFA 2020, NBER China Meeting 2020, MFA 2020, FMA 2020 etc.*
 - *Use the full sample of firms registered in China (40 million firms) to examine the network effect on firm growth, use centrality decomposition and geographic diversification to identify the channel.*

4. with Y. Yang, “Networks and Business Cycles” (Job Market Paper, [preprint](#))
 - *Talk: Princeton, UPenn*
 - *Develop a dynamic general equilibrium model with innovation network, production network, general cross-sectional shocks, and recursive preference.*
 - *Use a high dimensional dynamic state-space model (with EM algorithm) to estimate the underlying networks and shocks.*
 - *Show (in theory and data) that a combination of innovation network, production network, and cross-sectional shock can well predict economic recovery speed.*
 - *Big data: patent datasets (1911-2019) and input-output data (1951-2019).*
5. “Networks, Link Complexity, and Cross-Predictability” ([preprint](#))
 - **The Finalist of the Best Paper in Investment, FMA 2020.**
 - *Talk: FMA 2020, NSF 6h Annual Conference on Networks.*
 - *Data: patent datasets, Factset, Orbis and Who owns whom, I/E/B/S (analyst coverage), Thomas Reuters Institutional investors, Raven Packs News Analytics, Idiosyncratic volatility.*
 - *Show hedge fund managers can not address information incorporated in the weak but dense linkages, and propose a cross-momentum trading strategy with monthly excess alpha of [324 basis point](#).*
 - *Detailed analyses show this excess alpha is not driven by the traditional explanations: limited investor attention, arbitrage cost from idiosyncratic volatility, risk exposure, or earning announcement effect.*
6. with F. Allen, J. Cai, X. Gu, J. Qian, L. Zhao, “Centralization or Decentralization ? the Evolution of State Ownership in China.” ([preprint](#))
 - **XiYue Best Paper Award, China International Conference for Finance**
 - *Talk: CICF 2021, UPENN 2020, Wharton China Symposium*
7. “Networks, Long Run Risk, and Asset Pricing”. (*Draft available*)
 - *Develop a dynamic general equilibrium model incorporating innovation network, production network, general cross-sectional shocks, and recursive preference.*
 - *Show the low-rank structure of the innovation network and time-varying heterogeneous shock yields a large variation in the volatility of pricing kernel.*
 - *Show this channel explains the time-varying risk-premium puzzle and several cross-sectional puzzles in stock return.*
8. with J. Cai, H. Shen, D. Yang, L. Zhao, “Semi-supervised Learning in Networks”. (*Draft available*)
 - *Talk: Joint Statistical Meeting (American Statistical Association), 2021.*
 - *A predictive model with unobservable network.*
 - *L2 norm to boost the unobserved network and parameters in model.*
 - *Derive asymptotic distribution of parameters on predictive model and network.*
 - *Simulation: significantly boost estimation of parameters and the network.*
9. with J. Cai, L. Zhao “Reinforcement Learning in Dynamic Networks and Return Cross-Momentum”. (*In Progress*)
 - *Estimate dynamic networks and use the links to understand cross-momentum.*

- A reinforcing learning with the underlying network as dynamic state variables.
- Boosting Method 1: initialization using real network.
- Boosting Method 2: dynamic state variables as functions of the previous states, real observable networks (with noise), and historical stock returns.
- Boosting Method 3: quantile regressions in deep learning.
- Datasets: Five networks (self-constructed) and CRSP.

10. with J.Cai, L.Zhao, D.Yang, H.Shen, “Common v.s. Idiosyncratic Risk”. (*In Progress*)

INVITED TALKS

- 2021 China Financial Research Conference, China International Conference in Finance ($\times 3$), Asian Meeting for Econometric Society, Joint Statistical Meetings (American Statistical Association), American Finance Association Annual Conference, NSF 6th Annual Conference for Networks Economics($\times 4$), UChicago,Booth.
- 2020 NBER Chinese Economy Meeting*, Winter Meeting for Econometric Society, Financial Management Association Annual($\times 2$), Middle Western Finance Association ($\times 2$), American Economic Association Annual ($\times 2$)
- 2019 Summer Meeting Econometric Society, IMF($\times 2$)*, Jane Street PhD Symposium, Asian Meeting of Econometric Society, Bank of Finland*, Penn-Wharton-GSM, Penn Econ(Macro Lunch), Penn Econ(Micro Theory Lunch), Penn Econ (Micro Lunch),
- 2018 PKU, Penn Econ (Econometric Lunch), IMF*, Wharton(MBA Talk)
- 2016 Annual American Economic Association, Alibaba, NBER-CCER Conference*.

COURSES

Deep Learning in Theory, Optimization in Machine Learning, Non-Parametric & Machining Learning, Data Mining, Econometrics I, Bayesian Econometrics II, Econometrics IV, Continuous Time Asset Pricing, Asset Pricing, Empirical Methodology of Asset Pricing, Empirical Corporate Finance, Probability Theory, Stochastic Process (I, II), Measure Theory, Real Analysis, Financial Market and Macro Finance.

PH.D. ADVISORS

Rakesh Vohra (University Professor of Economics, ESE, and CIS), Linda H. Zhao (Professor of Statistics)