DINGYI LI

137 Reservoir Ave, Ithaca, NY 14850

(607) 262-4763 ◇ dl922@cornell.edu ◇ dingyili93.github.io

EDUCATION

- **Cornell University**
  Ph.D. in Applied Economics and Management   *2019-2026 (Expected)*

- **Renmin University of China**
  B.A. in Economics   *2012-2016*
  B.S. in Mathematics   *2012-2016*

- **Field of interest:** Econometrics, Industrial Organization

WORKING PAPERS

- "Identification and Estimation of Finite Mixtures of Multinomial Logit Models"

  Abstract: Finite mixtures of multinomial logit models can be used to capture consumer choice heterogeneity across multiple markets when only aggregate consumer choices per market are available. A motivating example is a nested logit where the composition of each mixture component (each nest of alternatives) is unknown a priori. We show that in order to identify these models, it suffices to require that each mixture component includes at least two component-exclusive alternatives. We refer to our assumption as the *pure-alternatives* condition, and we argue it is a natural extension of the *anchor-word* assumption used commonly in nonnegative matrix factorization problems in machine learning. Our identification result enables a consistent two-step estimator as the number of consumers, markets, and alternatives grow large. Applying this framework to the U.S. vehicle market, we find that consumer heterogeneity does not yield substitution patterns between electric and internal combustion engine vehicles, suggesting consumer segments are distinctly aligned with specific vehicle types without crossover substitution.

- "On the Testability of Anchor Words in Topic Models" (with Simon Freyaldenhoven, Shikun Ke, and José Luis Montiel Olea). *Revise & Resubmit at Quantitative Economics.*

  Abstract: Topic models are a simple and popular tool for the statistical analysis of textual data. Their identification and estimation is typically enabled by assuming the existence of *anchor words*; that is, words that are exclusive to specific topics. In this paper we show that the existence of anchor words is statistically testable: There exists a hypothesis test with correct size that has nontrivial power. This means that the anchor-words assumption cannot be viewed simply as a convenient normalization. Central to our results is a simple characterization of when a column-stochastic matrix with known nonnegative rank admits a *separable* factorization. We test for the existence of anchor words in two different data sets derived from monetary policy discussions in the Federal Reserve and reject the null hypothesis that anchor words exist in one of them.

- "Systemic Risk, FOMC Statements, and Monetary Policy Shocks: A New Topic Model to Associate Text with Metadata" (with Shawn Mankad).

  Abstract: In this research paper, we investigate the regulations guiding monetary policy communications through the development of a novel machine learning method called the Cluster Sentence Structural Topic Model (CSSTM). Our approach incorporates covariates in the data generation process and accounts for the correlation of sentences within each document by utilizing the equilibrium of sentences' topics. In the estimation process, we sort the equilibrium in the M step. Our method outperforms the Latent Dirichlet Allocation (LDA) and the Structural Topic Model (STM) by increasing the held-out likelihood by 20 percent and 10 percent. Using our method, we analyze FOMC statements and observe that the Fed places more emphasis on inflation expectations as opposed to current rates. According to our results, FOMC statements rely more on production instead of consumption. More importantly, we find that monetary policy communication started to consider systemic risk shortly after the 2007 financial crisis. By our method, we are able to decompose monetary policy shocks. The new measure has large and significant effects on systemic risk.

- "Pollution Avoidance and Willingness-to-Pay: Evidence from Travel Mode Choice in Beijing" (with Shanjun Li and C.-Y. Cynthia Lin Lawell).

Abstract: We estimate the short-term willingness-to-pay (WTP) to avoid air pollution by developing a model to capture the trade-offs between avoidance behavior and its costs. In particular, we use fine-scale travel survey data in Beijing to model the trade-offs between indoor and outdoor travel modes for compulsory work trips during highly polluted hours. Our model indicates that the short-term WTP, which we estimate to be 0.00223 dollars per hour to avoid 1 $\mu g/m^3$ of ambient fine particles ($PM_{2.5}$), forms the lower bound for the long-term WTP, which is around 11.536 dollars per year to avoid 1 $\mu g/m^3$ $PM_{2.5}$. Our estimation strategy uses a machine learning IV method in a high-dimensional econometrics setting. We find that longer potential exposure to air pollution prevents people from walking and cycling. People older than 55 years old, who are more vulnerable to pollution and thus more likely to avoid pollution, have a 28% higher WTP than the young. Likewise, richer people, who value their health more, are willing to avoid a unit of pollution with 36% more cost. Finally, we find evidence that information affects the behavioral adjustment: people start to reduce their exposure to the toxic air only after extensive media coverage of air pollution.

WORK IN PROGRESS

- "Weak Sparse Models and Methods for Instrumental Variables".

TEACHING EXPERIENCE

- Teaching Assistant for AEM 2300 International Trade and Finance, Cornell University   *2022-2024 Spring*
- Teaching Assistant for AEM 3310 Introduction to Business Regulation, Cornell University   *2023 Fall*
- Teaching Assistant for AEM 4110 Introduction to Econometrics (4.9/5.0), Cornell University   *2021 Fall*
- Teaching Assistant for AEM 6120 Applied Econometrics (4.4/5.0), Cornell University   *2020 Fall*

AWARDS

- Graduate Research Fellowship, Cornell University   *2024*
- Ashley Graduate Fellowship, Cornell University   *2023*

  Research Travel Grant, Cornell University

- Transportation Networks and Smart Mobility Scholarship, Massachusetts Institute of Technology   *2022*

  Research Travel Grant, Cornell University

- Edward and Janet Heslop Fellowship, Cornell University   *2021-2022*
- Academic Excellence Scholarship, Renmin University of China   *2012-2015*

PRESENTATIONS (*: CO-AUTHOR PRESENTS)

- New York Camp Econometrics XVIII, Cornell University, North American Summer Meeting of the Econometric Society, Hong Kong University of Science and Technology, International Association for Applied Econometrics (Xiamen and Thessaloniki), University of Chicago*, Econometric Society Interdisciplinary Frontiers Economics and AI+ML Meeting, Federal Reserve System Econometrics Meeting*   *2024*
- University of Cologne*, Cornell University, Asia Meeting of the Econometric Society, University of Chicago*, University of California Berkeley, University of Texas at Austin*   *2023*
- North American Summer Meeting of the Econometric Society, Philadelphia FED Conference on Frontiers in Machine Learning and Economics: Methods and Applications*   *2022*
- World Conference of Spatial Econometrics Association   *2021*

REFERENCES

José Luis Montiel Olea (Chair)
Associate Professor
Department of Economics
Cornell University
Phone: +1 (617) 821-1336
montiel.olea@gmail.com

Florentina Bunea
Professor
Department of Statistics and Data Science
Cornell University
Phone: +1 (607) 255-8449
fb238@cornell.edu

Francesca Molinari
H. T. Warshow and Robert Irving Warshow Professor
Department of Economics
Cornell University
Phone: +1 (607) 255-6367
fm72@cornell.edu