

THOMAS WIEMANN

wiemann@uchicago.edu – thomaswiemann.com – (312) 684-3417

Education

University of Chicago , Ph.D. Economics	2019–present
University of Oxford , M.Sc. Statistical Science	2018–2019
Erasmus University Rotterdam , B.Sc. Econometrics & Operations Research	2014–2018
Erasmus University Rotterdam , B.Sc. Economics & Business Economics	2014–2018

Working Papers

Optimal Categorical Instrumental Variables

Revise and Resubmit, Journal of Business & Economic Statistics

Abstract: This paper discusses estimation with a categorical instrumental variable in settings with potentially few observations per category. The proposed categorical instrumental variable estimator (CIV) leverages a regularization assumption that implies existence of a latent categorical variable with fixed finite support achieving the same first stage fit as the observed instrument. In asymptotic regimes that allow the number of observations per category to grow at arbitrary small polynomial rate with the sample size, I show that when the cardinality of the support of the optimal instrument is known, CIV is root- n asymptotically normal, achieves the same asymptotic variance as the oracle IV estimator that presumes knowledge of the optimal instrument, and is semiparametrically efficient under homoskedasticity. Under-specifying the number of support points reduces efficiency but maintains asymptotic normality. In an application that leverages judge fixed effects as instruments, CIV compares favorably to commonly used jackknife-based instrumental variable estimators.

A Practitioner’s Guide to Double/Debiased Machine Learning

with Achim Ahrens, Victor Chernozhukov, Christian Hansen, Damian Kozbur, Mark Schaffer
Revise and Resubmit, Journal of Economic Literature

Abstract: This paper provides a practical introduction to Double/Debiased Machine Learning (DML). DML provides a general approach to performing inference about a target parameter in the presence of nuisance parameters. The aim of DML is to reduce the impact of nuisance parameter estimation on estimators of the parameter of interest. We describe DML and its two essential components: Neyman orthogonality and cross-fitting. We highlight that DML reduces functional form dependence and accommodates the use of complex data types, such as text data. We illustrate its application through three empirical examples that demonstrate DML’s applicability in cross-sectional and panel settings.

Demand Estimation with Finitely Many Consumers

with Jonas Lieber

Abstract: Although market shares are frequently estimated via averages of finitely many consumer choices, commonly applied methods for demand estimation are not robust to estimation error in these shares. While non-negligible estimation error in market shares always introduces bias in the demand parameter estimators, the issue becomes most salient when estimated market shares are zero. In the presence of zero shares, widely applied estimators of the random coefficient logit model cannot be computed without ad-hoc data manipulations. This paper proposes a new estimator of demand parameters for settings with endogenous prices and estimated market shares that is robust to zero-valued market shares. The estimator generalizes the constrained optimization program of Dubé et al. (2012) with probabilistic bounds on the estimation error in market shares. We show consistency as the number of

markets T grows sufficiently slowly relative to the number of consumers n such that $\log(T)/n \rightarrow 0$, and provide confidence intervals under the same regime. Simulations suggest improved finite sample properties of the proposed estimator to conventional alternatives.

Effects of Health Care Policy Uncertainty on Households' Portfolio Choice

with Robin L. Lumsdaine

Abstract: This paper develops a nonparametric identification approach for causal effects of an endogenous macroeconomic variable on microeconomic outcomes. The key assumption is the existence of an exogenous variable that shifts responsiveness to the variable of interest without shifting responsiveness to other macroeconomic time series. We apply the approach to study the effect of health care policy uncertainty (HCPU) on households' portfolio choice using health shocks to capture cross-sectional heterogeneity. Under the additional assumption of risk averse agents, our approach provides an informative bound on the average causal effect of HCPU. The empirical results highlight HCPU as an important determinant of households' financial behavior, and showcase substantial heterogeneity in HCPU effects across varying unexpected changes to health.

Publications & Forthcoming

Model Averaging and Double Machine Learning

with Achim Ahrens, Christian Hansen, Mark Schaffer

Journal of Applied Econometrics, forthcoming.

Abstract: This paper discusses pairing double/debiased machine learning (DDML) with stacking, a model averaging method for combining multiple candidate learners, to estimate structural parameters. We introduce two new stacking approaches for DDML: short-stacking exploits the cross-fitting step of DDML to substantially reduce the computational burden and pooled stacking enforces common stacking weights over cross-fitting folds. Using calibrated simulation studies and two applications estimating gender gaps in citations and wages, we show that DDML with stacking is more robust to partially unknown functional forms than common alternative approaches based on single pre-selected learners. We provide Stata and R software implementing our proposals.

ddml: Double/debiased machine learning in Stata

with Achim Ahrens, Christian Hansen, Mark Schaffer

Stata Journal, 2024, 24(1): 3-45.

Abstract: We introduce the package ddml for Double/Debiased Machine Learning (DDML) in Stata. Estimators of causal parameters for five different econometric models are supported, allowing for flexible estimation of causal effects of endogenous variables in settings with unknown functional forms and/or many exogenous variables. ddml is compatible with many existing supervised machine learning programs in Stata. We recommend using DDML in combination with stacking estimation which combines multiple machine learners into a final predictor. We provide Monte Carlo evidence to support our recommendation.

Work in Progress

Targeting Profits in Observational Consumer Panels with Andrew Bai, Sanjog Misra

Bayesian ML for Estimation of Observed and Unobserved Consumer Heterogeneity

Sign Reversing and Sign Preserving Control Function Estimands

Awards, Scholarships, and Grants

Robert Lucas Jr. Fellowship, University of Chicago	2024–2025
CV Starr Fellowship, University of Chicago	2024–2025
J. Lawrence Laughlin Fellowship, University of Chicago	2024–2025
Graduate Fellowship, University of Chicago	2019–2024
ERP Fellowship, German Ministry for Economic Affairs and Energy	2019–2020
Fellowship, German Academic Scholarship Foundation	2015–2019

Teaching Experience

Applied Bayesian Econometrics (PhD)	TA for Prof. Sanjog Misra	Winter 2025
Econometrics (Undergraduate)	College Lecturer	Spring 2022
Optimization-Conscious Econometrics (PhD)	TA for Prof. Guillaume Pouliot	Winter 2022
Applied Microeconometrics (PhD)	TA for Prof. Alexander Torgovitsky	Fall 2021
Econometrics (Undergraduate)	TA for Prof. Max Tabord-Meehan	Spring 2021

Research Experience and Other Employment

Research Assistant for Prof. Christian Hansen, Booth School of Business, Chicago	2020–2024
Economist Intern, Consumer Behavior Analytics Team, Amazon, Seattle	2022
Research Assistant for Prof. Philip Hans Franses, Econometric Institute, Rotterdam	2018
Research Assistant for Prof. Zareh Asatryan, ZEW, Mannheim	2017
Research Assistant for Prof. Sacha Kapoor, Erasmus University Rotterdam, Rotterdam	2016

Professional Experience

Co-founder and Organizer of the Comp. Methods in Econ. Workshop, University of Chicago	2021–2023
Graduate Student Liaison, University of Chicago	2020–2023
Undergraduate Student Liaison, Erasmus University Rotterdam	2015–2017

Conferences	2024: <i>North American Winter Meeting of the Econometric Society (San Antonio)</i>
	2023: Royal Economic Society (Glasgow), Optimization Conscious Econometrics Conference (Chicago), North American Summer Meeting of the Econometric Society (UCLA), International Association for Applied Econometrics (Oslo)
	2022: Summer Institute for Applied Artificial Intelligence (Chicago)
	2019: Society for Financial Econometrics (Shanghai), International Association for Applied Econometrics (Nicosia), Stanford Institute for Theoretical Economics Workshop on Macroeconomic Uncertainty (Stanford)

Refereeing Activity	<i>Journal of Econometrics</i>
----------------------------	--------------------------------

Software

ddml: Double/debiased machine learning in Stata

ddml: Double/debiased machine learning in R

kcmeans: Conditional expectation function estimation with K-conditional-means

civ: Categorical instrumental variables

Additional Information

Programming Skills R, C++, Julia, Python, PySpark, SQL, MATLAB, Stata

Languages German (Native), English (Fluent)

Citizenship German