EDVARD BAKHITOV
www.edbakhitov.com
bakhitov@sas.upenn.edu

UNIVERSITY OF PENNSYLVANIA

Placement Director: Holger Sieg	HOLGERS@ECON.UPENN.EDU	215-898-7194
Placement Director: Iourii Manovski	MANOVSKI@ECON.UPENN.EDU	215-898-6880
Graduate Student Coordinator: Gina Conway	GNC@SAS.UPENN.EDU	215-898-5691

Office Contact Information
The Ronald O. Perelman Center for Political
Science and Economics, Room 635
133 South 36th Street
Philadelphia, PA, 19104
Phone: 215-450-0973

Personal Information:
Date of Birth: June 5th, 1992
Citizenship: Russia (F-1 Visa)

Undergraduate Studies:
B.S., Economics, Higher School of Economics, Russia, 2013

Masters Level Work:
M.S., Economics, Higher School of Economics, Russia, 2015
M. S., Economics and Finance, University of Luxembourg, Luxembourg, 2015

Graduate Studies:
University of Pennsylvania, 2016 to present
Thesis Title: "*Machine Learning under Endogeneity*"
Expected Completion Date: May 2022

Thesis Committee and References:

Professor Amit Gandhi (Advisor)
Department of Economics
University of Pennsylvania
133 South 36th Street, Suite 622,
Philadelphia, PA, 19104
Phone: 215-898-7409
E-mail: akgandhi@sas.upenn.edu

Professor Xu Cheng
Department of Economics
University of Pennsylvania
133 South 36th Street, Suite 620,
Philadelphia, PA, 19104
Phone: 215-898-6775
E-mail: xucheng@econ.upenn.edu

Professor Karun Adusumilli
Department of Economics
University of Pennsylvania
133 South 36th Street, Suite 631,
Philadelphia, PA, 19104
Phone: 215-898-7676
E-mail: akarun@sas.upenn.edu

Professor Francis X. Diebold
Department of Economics
University of Pennsylvania
133 South 36th Street, Suite 607,
Philadelphia, PA, 19104
Phone: 215-898-1507
E-mail: fdiebold@sas.upenn.edu

Teaching and Research Fields:

Econometrics, Machine Learning, Industrial Organization

Teaching Experience:

Fall 2017	Statistics for Economists (Undergraduate), Teaching assistant for Doctor Suleyman Ozmucur
Spring 2018	Macroeconomic Theory (Undergraduate), Teaching assistant for Professor Dirk Krueger
Fall 2018	Introduction to Econometrics (Undergraduate), Teaching assistant for Professor Francis Diebold
Spring 2019	Market Design (Undergraduate), Teaching assistant for Professor Amit Gandhi
Spring 2020	Market Design (Undergraduate), Teaching assistant for Professor Amit Gandhi
Fall 2020	Statistics for Economists (Undergraduate), Teaching assistant for Professor Karun Adusumilli
Spring 2021	Statistics for Economists (Undergraduate), Teaching assistant for Professor Wayne Gao

Research Experience and Other Employment:

2018 – 2020	University of Pennsylvania, Research assistant for Professor Amit Gandhi, Professor Aviv Nevo and Professor Jing Tao. Project: *Flexible Estimation of Discrete Choice Models using Aggregate Data.*

Professional Activities:

2019	EEA-ESWM
2021	UAI – Advances in Causal Inference Workshop (poster session)
	ICML – ML4Data Workshop (poster session)
	ICML – Neglected Assumptions in Causal Inference Workshop (poster session)

Honors, Scholarships, and Fellowships:

2016 – 2021	University of Pennsylvania Fellowship
2015	Best Master Thesis Prize 2015 in Economics by the Economist Club Luxembourg and the University of Luxembourg's Faculty of Law, Economics, and Finance
2014	Luxembourg Ministry of Culture Scholarship

Research Papers:

"*Automatic Debiased Machine Learning in Presence of Endogeneity*" **(Job Market Paper)**

Recent advances in machine learning literature provide a series of new algorithms that both address endogeneity and can be applied in high-dimensional environments, we call them MLIV. This paper introduces an approach for performing valid asymptotic inference on regular functionals of MLIV estimators. The approach is based on construction of an orthogonal moment function that has a zero derivative with respect to the MLIV estimator. The debiasing is automatic in the sense that it only depends on the form of the identifying moment function but not on the form of the bias correction term. We derive a convergence rate for the penalized GMM estimator of the bias correction term. We also give conditions for root-n consistency and asymptotic normality of the debiased MLIV estimator of the functional of interest. Overall, the approach allows for a large variety of MLIV estimators as long as they satisfy mild convergence rate conditions. We apply our procedure to estimate the conditional demand derivative within the nonparametric demand for differentiated goods framework. Using both simulated and real data, we demonstrate that our debiased estimates have significantly reduced bias and close to the nominal level coverage, while the plug-in estimates perform poorly.

"*Causal Gradient Boosting: Boosted Instrumental Variables Regression*" (with Amandeep Singh)

Recent advances in the literature have demonstrated that standard supervised learning algorithms are ill-suited for problems with endogenous explanatory variables. To correct for the endogeneity bias, many variants of nonparameteric instrumental variable regression methods have been developed. In this paper, we propose an alternative algorithm called boostIV that builds on the traditional gradient boosting algorithm and corrects for the endogeneity bias. The algorithm is very intuitive and resembles an iterative version of the standard 2SLS estimator. Moreover, our approach is data driven, meaning that the researcher does not have to make a stance on neither the form of the target function approximation nor the choice of instruments. We demonstrate that our estimator is consistent under mild conditions. We carry out extensive Monte Carlo simulations to demonstrate the finite sample performance of our algorithm compared to other recently developed methods. We show that boostIV is at worst on par with the existing methods and on average significantly outperforms them.

"*Frequentist Shrinkage under Inequality Constraints*"

This paper shows how to shrink extremum estimators towards inequality constraints motivated by economic theory. We propose an Inequality Constrained Shrinkage Estimator (ICSE) which takes the form of a weighted average between the unconstrained and inequality constrained estimators with the data dependent weight. The weight drives both the direction and degree of shrinkage. We use a local asymptotic framework to derive the asymptotic distribution and risk of the ICSE. We provide conditions under which the asymptotic risk of the ICSE is strictly less than that of the unrestricted extremum estimator. The degree of shrinkage cannot be consistently estimated under the local asymptotic framework. To address this issue, we propose a feasible plug-in estimator and investigate its finite sample behavior. We also apply our framework to gasoline demand estimation under the Slutsky restriction.

Research Papers in Progress:

"*Deep Causal Inequalities: Demand Estimation using Individual-Level Data*" (with Amandeep Singh and Jiding Zhang)

Modern machine learning algorithms can easily deal with unstructured data, however, recent literature has demonstrated that they do not perform well in presence of endogeneity in the explanatory variables. On the other hand, extant methods catered towards addressing endogeneity issues make strong parametric assumptions and, hence, are incapable of directly incorporating high-dimensional unstructured data. In this paper, we borrow from the literature on partial identification and propose the Deep Causal Inequalities (DeepCI) estimator that overcomes both these issues. Instead of relying on observed labels, the DeepCI estimator uses inferred moment inequalities from the observed behavior of agents in the data. This allows us to take care of endogeneity by differencing out unobservable product characteristics. We provide theoretical guarantees for our estimator and prove its consistency under very mild conditions. We demonstrate through extensive Monte Carlo simulations that our estimator outperforms standard supervised machine learning algorithms and existing partial identification methods. Finally, we apply DeepCI to the differentiated products demand estimation framework. The flexibility of the method allows for highly unstructured data like images, which we exploit in the empirical application based on the consumer-level car rental data from Hertz. Using the DeepCI estimator, we show how to estimate the importance of various car design features affecting consumer rental decisions.

"*Feature Selection in Differentiated Product Demand Models*" (with Amit Gandhi and Jing Tao)

We provide a novel approach to modeling substitution patterns in differentiated products demand models by arguing that consumer choices are driven by product "features" rather than raw product characteristics. In order to "featurize" a differentiated product demand the model, we apply recent innovations in machine learning to a nonlinear structural estimation problem. Our approach is based on treating the inverse demand function in a differentiated product market as the primitive of interest and exploit restrictions on the inverse demand imposed by the mixed logit structure to express the estimation problem as a partially linear regression with many endogenous variables. Since the true features are unknown to the econometrician, we propose to saturate the model with potential features and apply the Double Lasso estimator to perform selection of both relevant features and relevant instruments. By exploiting the implicit function theorem, we directly recover price elasticities from the inverse demand function estimates without estimating the distribution of random coefficients. A simulation study shows that the proposed estimation procedure performs well in finite samples.