Harvard Medical School
Curriculum Vitae

Date Prepared:	October 27, 2020
Name:	José R. Zubizarreta
Office Address:	Department of Health Care Policy Harvard Medical School Harvard University 180 Longwood Ave, Office 307-D Boston, MA 02115
Work Phone:	(617) 432-1266
Work Email:	zubizarreta@hcp.med.harvard.edu
Work Fax:	(617) 432-3503

Education:

2005	BS	Mathematical Engineering	Universidad Católica de Chile
2008	MA	Economics	Universidad de Chile
2011	MA	Statistics	The Wharton School, University of Pennsylvania
2013	PhD	Statistics	The Wharton School, University of Pennsylvania

Faculty Academic Appointments:

07/13-12/13	Instructor (Convertible)	Division of Decision, Risk, and Operations	Columbia Business School, Columbia University
10/13-06/17	Assistant Professor (by courtesy)	Department of Statistics	Faculty of Arts and Sciences, Columbia University
01/14-06/17	Faculty Affiliate	Data Science Institute	Columbia University
01/14-06/17	Assistant Professor	Division of Decision, Risk, and Operations	Columbia Business School
01/16-	Visiting Assistant Professor	Department of Pediatrics	School of Medicine, Universidad Católica de Chile
07/17-06/18	Assistant Professor	Department of Health Care Policy	Harvard Medical School, Harvard University

07/17-	Faculty Affiliate	Department of Statistics	Faculty or Arts and Sciences, Harvard University
07/18-	Associate Professor	Department of Health Care Policy	Harvard Medical School, Harvard University
03/20-	Faculty Affiliate	Harvard Data Science Initiative	Harvard University
06/20-	Associate Professor	Department of Biostatistics	Harvard School of Public Health, Harvard University

Other Professional Positions:

2005-2006	Coordinator of Voluntary Teachers	INFOCAP (NGO that provides labor training to low incomes workers; Chile)	*52 weeks per year*
2005-2006	Research Assistant	Universidad Católica de Chile	*12 hrs per week (52 weeks)*
2005-2006	Associate Researcher	Observatorio Social, Universidad Alberto Hurtado (Chile)	*12 hrs per week (52 weeks)*
2007-2008	Coordinator Area of Economic Studies	Observatorio Social, Universidad Alberto Hurtado (Chile)	*52 weeks per year*
2009	Research Assistant	Center for Promotion of Research Involving Innovative Statistical Methodology, NYU	*12 weeks per year*
2010-2013	Research Assistant	Department of Statistics, The Wharton School, University of Pennsylvania	*20 hrs per week (52 weeks)*

Major Administrative Leadership Positions:

Regional:

2016	Conference Organizer	Columbia University Causal Inference Conference on Point Exposures
2016	Conference Organizer	Columbia University Causal Inference Conference on Effect Heterogeneity
2017	Conference Organizer	Columbia University Causal Inference Conference on Longitudinal Studies

National:

2013	Assistant Program Director	MIT Media Lab Encuentros Conference

Committee Service:

Local:

2014-2017	PhD Admissions Committee	Columbia Business School
2015-2017	Faculty Computing Committee	Columbia Business School
2015-2017	Empowering Research Committee	Columbia Business School
2020-	Standing Committee on Health Policy	Faculty of Arts and Sciences, Harvard University

National:

2018	Reviewer	National Science Foundation (NSF), Methodology, Measurement, and Statistics (MMS) Program
2015, 2018, 2019	Thomas R. Ten Have Award Committee	Atlantic Causal Inference Conference
2018, 2019	Reviewer	Patient Centered Outcomes Research Institute (PCORI), Methods Panel

International:

2016	Session Chair	INFORMS Annual Meeting
2017, 2020	Student Awards Committee	International Conference in Health Policy Statistics
2017	Byar Award Committee	Biometrics Section of the American Statistical Association
2017, 2018	Paper Review Committee	Society for Research on Educational Effectiveness
2018	Reviewer	British Medical Research Council
2020	COVID-19 Metrics Task Force	Joint committee formed by Universidad Católica, Universidad de Chile and Universidad de Concepción in coordination with the Government of Chile to measure the COVID-19 pandemic

Professional Societies:

2011-	American Statistical Association	Member
2011-	International Biometric Society, Eastern North American Region	Member
2011-	Institute of Mathematical Statistics	Member

Editorial Activities:

Editorial Roles:

2018-	Associate Editor	Observational Studies
2019-	Associate Editor	Journal of Computational and Graphical Statistics
2020-	Associate Editor	Biometrics
2020-	Co-Editor	Handbook of Propensity Score Methods (Chapman & Hall)

Adhoc Reviewer:

American Journal of Epidemiology
American Journal of Political Science
American Journal of Psychiatry
Annals of Applied Statistics
Annals of Epidemiology
Annals of Operations Research
Annals of Statistics
Annals of Surgery
Biometrics
Biometrika
Biostatistics
BMC Medical Research Methodology
Computational Statistics and Data Analysis
Epidemiology
Epidemiologic Methods
Health and Services and Outcomes Research Methodology
Journal of Applied Econometrics
Journal of Causal Inference
Journal of Computational and Graphical Statistics
Journal of Educational and Behavioral Statistics
Journal of Machine Learning Research
Journal of Research on Educational Effectiveness
Journal of the American Statistical Association, Applications and Case Studies
Journal of the American Statistical Association, Theory and Methods
Journal of the Royal Statistical Society: Series A (Statistics in Society)
Journal of the Royal Statistical Society: Series B (Statistical Methodology)
Journal of the Royal Statistical Society: Series C (Applied Statistics)
Management Science
Observational Studies
Operations Research
PLOS ONE
Psychometrika
Review of Economics and Statistics

Statistica Sinica
Statistical Science
Statistical Methods in Medical Research
Statistics and Computing
Statistics and Public Policy
Statistics in Medicine
Trials
The American Statistician

Honors and Prizes:

2005	Maximum Distinction	Universidad Catolica de Chile	Mathematical Engineering
2008-2012	Fulbright Scholarship		
2009-2013	Graduate Fellow	The Wharton School, University of Pennsylvania	
2011	J. Parker Memorial Bursk Prize	Statistics Department, The Wharton School, University of Pennsylvania	For Excellence in Research
2011	Thomas R. Ten Have Memorial Award	Atlantic Causal Inference Conference	Award for "exceptionally creative or skillful research on causal inference"
2011	Student Paper Award	Health Policy Statistics Section, American Statistical Association	
2012	Deming Scholar Award	American Statistical Society/American Society for Quality	
2012	Young Investigator Award	Statistics in Epidemiology, American Statistical Association	
2012	President Gutmann Leadership Award	University of Pennsylvania	
2013	Student Paper Award	Social Statistics, Government Statistics, and Survey Research Methods Sections, American Statistical Association	
2014	Student Paper Award	International Biometric Society ENAR	
2014	Kenneth Rothman Prize	Epidemiology	For the best paper published in Epidemiology in 2013
2017	Initiative on Data Science Visiting Fellow	Booth School of Business, University of Chicago	
2017	Runner-up Pierskalla Best Paper Award	Health Applications Society, INFORMS	
2019	Runner-up Ralph Gomory Best Industry Studies Paper Award	Industry Studies Association	
2020	William Cochran Prize	Observational Studies	For the best paper published in Observational Studies in 2013-2020

Report of Funded and Unfunded Projects

Funding Information:

Past:

2008-2009 Educational Outcomes of the Children of the Poor: The Chilean Case
United Nations Development Plan
Principal Investigator
In 2006, one out of every five children lived below the poverty line in Chile. That same year, half of the children in the country had suffered poverty at least once during their childhood. This project asks the question, what is the impact of experiencing poverty during childhood on educational outcomes during early adulthood? Addressing this question can enlighten the mechanisms under which poverty is transmitted across generations and help to understand the forces that generate and perpetuate the inequality of the Chilean Society.

2015-2017 New Methods for Causal Inference in Randomized and Observational Studies
Alfred P. Sloan Foundation
Principal Investigator
Randomized experiments constitute the most reliable device for learning about the effects of treatments, policies or interventions on human subjects. Nonrandomized or observational studies are ubiquitous in the health and social sciences in part because harmful treatments cannot be imposed to individuals for experimentation. The goal of this research program is to develop new statistical methods that improve the design and analysis of both randomized experiments and observational studies of causal effects. The specific objectives of this project are: (i) to develop new statistical methods that improve (i.a.) the degree of control and (i.b.) the efficiency of randomized experiments, while providing a justified basis for statistical inference and (ii) to develop new, alternative statistical methods to those based on model-based estimates of the propensity score to (ii.a) better adjust for observed covariates and (ii.b) yield more stable estimates causal estimates, especially in longitudinal studies of treatment effects.

2016-2021 Health and Disability over the Life Course
NIH, NIA/R01AG056238
Project Co-Investigator (Maestas)
This project investigates the health and work capacity of individuals with moderate and severe health problems, examining how work capacity varies with characteristics of the disability and in relation to economic conditions over time, using administrative and survey data.

2016-2021 Disability Among Older Low-Skilled Workers
NIH, NIA/R01AG056239
Co-Investigator (Maestas)
This project investigates the health and work abilities of low-skilled, older individuals with moderate health problems, by education and work history, and in relation to economic conditions over time, using administrative and survey data.

2017-2022 Machine Learning for Health Outcomes and Quality of Care in Low-Income Populations
NIH/1DP2MD012722
Co-Investigator (Rose)

This project will develop a novel machine learning framework for the generalizability of experimental and quasi-experimental studies, providing population health scientists with robust methodology to assess the effects of health interventions and exposures. Health outcomes and quality of care in low- income populations lag behind other groups, and the impact of health insurance on these disparities among low-income individuals is currently unknown. A major goal of this proposal is to examine the role of insurance coverage on health outcomes in low-income populations with rigorous new tools in partially randomized data.

2018-2023 Alzheimer's Disease and Related Dementia Care within the Medicare Program
NIH /5R01AG062282
Co-Investigator (Hsu)
The current project will use a novel dataset that contains individual-level linkages between three data sources: 1) comprehensive Medicare claims from Parts A, B, and D; 2) EHR data from a large health system; and 3) a dementia research registry. The project also will use national traditional fee-for-service Medicare data and a P01-developed instrument on MA benchmarks to assess spillover effects within TM. We will address three specific aims within the overall population and among vulnerable subgroups, e.g., racial/ethnic minority or dual-eligible beneficiaries: Aim 1. Validation of Claims-based ADRD Disease and Severity Definitions: We will validate algorithms using claims data for identifying beneficiaries with ADRD of varying stages or levels of severity, combined with a gold-standard definition based on EHR chart reviews. Aim 2. MA Spillover Effects on ADRD Diagnostic Quality in TM: We will assess the impact of MA penetration (using an IV based on changes in MA benchmarks) on guideline concordant or discordant testing. Aim 3. MA Spillover Effects on ADRD Treatment Quality in TM: We will assess the impact of MA penetration on guideline concordant or discordant treatments.

Current:

2017-2020 Improving Feasibility and Efficiency of Quality Measurement in Oncology Practices
2020-2023 Laura and John Arnold Foundation / 20-04402
Co-Investigator (Chernew)
This project will first estimate the extent to which claims-based measures of quality for practices (with and without clinical information about cancer stage and other tumor characteristics from registry data) correlate with practice-reported measures of oncology care quality. Next, for measures of quality for which supplemental clinical data are needed, assess the extent to which adaptive quality measurement strategies that focus primarily on identifying the providers who are delivering the highest- and lowest-quality care would be a more feasible and efficient strategy for assessing (and rewarding) quality of care delivered by oncology providers than measuring quality for all oncology practices.

2018-2021 New Statistical Methods for Causal Inference in the Social Sciences and Public Policy
Alfred P. Sloan Foundation/G-2018-10118 $284,377
Principal Investigator
This project proposes new statistical methods for causal inference in the social sciences and public policy. The specific objectives of this project are: (i) to develop new statistical methods that flexibly adjust for covariates and yield more stable causal estimates in observational studies with (i.a) instrumental variables and (i.b) multivalued treatments; (i.c) to study the formal properties of these and related methods; (ii) to

develop a new framework for the design and analysis of observational studies with discontinuities that facilitates identification, estimation, and generalization beyond the cutoff of the running variable; and (iii) to develop new methods that improve the degree of control (covariate balance) and statistical efficiency of randomized experiments and enhance their generalizability.

2020-2023 Leveraging EHR Data to Evaluate Key Treatment Decisions to Prevent Suicide-Related Behaviors
NIH/NIMH1/R01 MH121478-01
Co-Investigator (Kessler)
This project aims to develop precision treatment rules for primary care physicians trying to develop a treatment plan for patients seeking treatment for common mental disorders and suicide prevention coordinators trying to develop a treatment plan for a patient who just made a nonfatal suicide attempt with the goal of developing a plan that will minimize prevalence of suicide-related behaviors (either suicide deaths or nonfatal attempts) over the next 12 months. The study will carry out two prospective observational studies using Electronic Health Records to evaluate effects of key treatment decisions on suicide-related behaviors over the next 12 months.

2020-2022 Causal Inference with Complex Treatment Regimes: Design, Identification, Estimation, and Heterogeneity
Alfred P. Sloan Foundation / G-2020-13946
Co-Investigator (Dominici)
As real-world evidence becomes more prominent across the spectrum of social sciences, and particularly in economics, new challenges emerge for randomized and observational studies. Such challenges pertain both to the design phase of a causal analysis as well as to the estimands of interest. In this project, we will address three main issues of critical importance in social sciences: (i) potential failure of randomized control trials (RCT) to balance the pre-treatment variables, especially in high- dimensional settings; (ii) estimation of spillover effects (effects that arise when the outcome of one unit is affected by the treatment received by the other units) in settings where individuals interact with one another; (iii) interpretable discovery and inference on the heterogeneity of causal effects in scenarios with complex treatment regimes. Managing these issues is critical to advancing the field of causal inference.

2020-2023 Towards a New Generation of Matching Methods for Comparative Effectiveness Research
Patients Centered Outcomes Initiative (PCORI)/ME-2019C1-16172 $749,055
Principal Investigator
The overarching goal of this project is to develop a new generation of matching methods that can be used to directly and flexibly balance baseline covariates in comparative effectiveness research (CER) with big, complex, and rich observational data sets. The proposed methods will improve substantially over existing propensity score matching (PSM) and related matching approaches in three ways, all of them building on previous research by the PI that developed direct matching methods to replace the indirect methods used in PSM: Aim 1 will focus on matching for big data by leveraging recent advancements in computation and optimization to (Aim 1.1) scale direct balancing matching methods to large data sources, (Aim 1.2) target causal parameters for specific populations of interest, and (Aim 1.3) devise a data-driven algorithm that will allow the investigator to make explicit bias-variance tradeoffs in approximate covariate balancing. Aim 2 will focus on matching for heterogeneous treatment effects (HTE) and

personalized medicine (PM) by developing a new matching strategy to (Aim 2.1) balance covariates for specific treatment-covariate subgroups and test for HTE, (Aim 2.2) find the largest matched sample that represents a particular patient of interest to evaluate personalized treatments, and (Aim 2.3) establish the large sample properties of matching methods that directly balance covariates and devise the first formalized doubly robust matching estimator. Aim 3 will focus on matching for high-dimensional treatments by extending the above methods to (Aim 3.1) estimate the effects of multi-valued (non-binary) treatments, allowing individual health care providers to be considered "treatments," and thereby to (Aim 3.2) make possible to assess the quality of health care providers for given case-mixes of patients. In each of these aims, we will evaluate the performance of the proposed methods both in simulated and in empirical exemplar data sets from the Veterans Health Administration (VHA). The exemplar data sets will allow us to illustrate the value of the new methods by addressing questions in mental health research that are of independent interest to physicians, policymakers, and stakeholders. Dissemination is a critical objective, with our final aim (Aim 4) to develop open- and easy-to-use software, case study vignettes, and tutorials to make this new generation of matching methods widely available to practitioners in CER and PCOR.

2019-2023 The Impact of Telestroke on Patterns of Care and Long-Term Outcomes
NIH/NINDS / R01NS111952
Co-Investigator (Mehrotra)
Many patients with a stroke, in particular those in rural communities, receive care at an emergency department that does not have a stroke expert and therefore they may not receive life-saving reperfusion. Telestroke is one potential solution; with telestroke a stroke expert who is physically far away guides a local physician through the decision-making process via videoconference. In this project, our goal is to understand the impact of telestroke on where patients get care, likelihood of dying, disability, and the experience of front-line physicians and nurses.

Projects Submitted for Funding:

Report of Local Teaching and Training

Teaching of Students in Courses:

2007-2008	Applied Econometrics Undergraduate Students	Universidad Catolica de Chile, Department of Industrial and Systems Engineering 1.5-hr sessions 2x per week for 12 weeks
2014	Introduction to Econometrics and Statistical Inference PhD Students	Columbia Business School, Division of Decision, Risk, and Operations 1.5-hr sessions 2x per week for 12 weeks
2014-2016	Causal Inference PhD Students	Columbia Business School, Division of Decision, Risk, and Operations 1.5-hr sessions 2x per week for 12 weeks
2015-2016	Managerial Statistics MBA Students	Columbia Business School, Division of Decision, Risk, and Operations 1.5-hr sessions 3x per week for 6 weeks (taught 2 sections)
2017-2019	Methods Seminar PhD Students	Harvard Medical School, Department of Health Care Policy 1-hr sessions 1x month for 4 months

2017-	Health Policy Statistics Reading Course PhD Students	Harvard Medical School, Department of Health Care Policy 1.5-hr session 1x month for 4 months
2018-2019	Design of Experimental and Non-experimental Studies (DENS) Reading Group PhD Students	Harvard Medical School, Department of Health Care Policy; Faculty of Arts and Sciences, Department of Statistics 1-hr session 1x per week for 9 months
2018-	Causal Inference Reading Group PhD Students	Harvard Medical School, Faculty of Arts and Sciences, and Graduate School of Education 1-hr session 2x per month for 9 months
2019-	Essentials of the Profession Medical Students	Harvard Medical School 1-hr tutorial 2x per week for 2.5 weeks
2019-	Stat-397 Design of Experimental and Non-experimental Studies PhD seminar course	Faculty of Arts and Sciences, Department of Statistics 1.5-hr session 4x per month fall semester
2020	HKS API-115 / Economics 2115 / HBS-4175 Econometric Methods for Applied Research II	Harvard Kennedy School 1.5-hr session 2x for 9 months
2020-	2000B/SUP-958/HPM-246 Health Policy Research Designs and Methods	Faculty of Arts and Sciences, Health Policy PhD program 3-hr session 1x for 9 months
2020-	Stat-293/Stat-393 Design of Experimental and Non-experimental Studies Graduate course	Faculty of Arts and Sciences, Department of Statistics 3-hr session 4x per month fall and spring semesters

Research Supervision:

2017-	Supervision of four PhD candidates from FAS and HSPH	Individual meetings 1 hour per week each
2017-	Supervision of two MA candidates from FAS and HSPH	Individual meetings 1 hour per week each

Formally Mentored Harvard Medical, Dental and Graduate Students:

2017-2018 Christopher Hase, MA Candidate in Statistics, Harvard University
Published manuscript in *Statistics in Medicine*

2017-2019 Reagan Moser, PhD Candidate in Statistics, Harvard University
Published manuscript in the *Handbook of Research Methods in Clinical Psychology*
One manuscript in preparation

2017- Ambarish Chattopadhyay, PhD Candidate in Statistics, Harvard University
Advisor
Published manuscript in *Statistics in Medicine*
Three manuscripts in preparation

2017-2019 Juan Díaz, PhD Candidate in Statistics, Harvard University
Advisor
One submitted manuscript

One manuscript in preparation
Now: Assistant Professor, Universidad de Chile

2018-2019 Zacharias Branson, PhD Candidate in Statistics, Harvard University
Dissertation Committee Member
One manuscript in preparation
Now: Assistant Professor, Carnegie Mellon University

2018-2019 Debmalya Mandal, PhD Candidate in Computer Science, Harvard University
Dissertation Committee Member
One manuscript in preparation
Now: Postdoctoral Fellow, Columbia University

2018- Xiao Wu, PhD Candidate in Biostatistics, Harvard University
Dissertation Committee Member

2019- Bijan Niknam, PhD Candidate in Health Policy (Methods for Policy Research), Harvard University
Advisor
Two manuscripts in preparation

2020- Eric Cohn, PhD Candidate in Biostatistics, Harvard University
TBD
Two manuscripts in preparation

2020- Eric Dunipace, PhD Candidate in Biostatistics, Harvard University
Dissertation Committee Member
Two manuscripts in preparation

2020- Chen Lu, PhD Candidate in Statistics, Harvard University
TBD
One manuscript in preparation

2020- Yige Li, PhD Candidate in Biostatistics, Harvard University
Advisor
Four manuscripts in preparation
Three statistical software packages in preparation

2020- Kwangho Kim, Seidman Postdoctoral Fellow, Harvard University
Mentor

2020- Shasha Han, Postdoctoral Fellow, Harvard University
Mentor

2020- Yi Zhang, PhD Candidate in Statistics, Harvard University
Advisor
One manuscript in preparation

Other Mentored Trainees and Faculty:

2014-2015 Nikhil Bhat, PhD in Decision, Risk, and Operations, Columbia University
Dissertation Committee Member
Now: Software Engineer, Google Research

2014-2016 Cinar Kilcioglu, PhD in Decision, Risk, and Operations, Columbia University
Co-Advisor
Published manuscript in the *Annals of Applied Statistics*
Now: Senior Data Scientist, Uber Research

2014-2016 Zach Shahn, PhD in Statistics, Columbia University
Oral Exam and Dissertation Committee Member
One manuscript in preparation
Now: Postdoctoral Fellow, Harvard School of Public Health

2014-2017 Wengi Hu, PhD in Decision, Risk, and Operations, Columbia University
Co-Advisor
Published manuscript in *Manufacturing & Service Operations Management*
**Runner-up for the Pierskalla Best Paper Award, 2017 (awarded by the Health Applications Society of INFORMS for research excellence in the field of health care management science)*
**Finalist for the MSOM Student Paper Award, 2017 (awarded by the MSOM Society of INFORMS for papers judged to be the best in the field of operations management)*
Now: Data Scientist, Uber Research

2014-2017 Maria Resa, PhD in Statistics, Columbia University
Co-Advisor
Published manuscript in *Statistics in Medicine*
Published manuscript in the *Journal of the Royal Statistical Society – Series A*
**Winner of the 2017 Student Paper Award of the Social Statistics, Government Statistics and Survey Research Methods Sections of the American Statistical Association*
One submitted manuscripts
Now: Data Scientist, Pfizer Research

2014-2017 Susanna Makela, PhD in Statistics, Columbia University
Oral Exam and Dissertation Committee Member
Now: Data Scientist, Google Research

2015-2018 Giancarlo Visconti, PhD in Political Science, Columbia University
Dissertation Committee Member
Published manuscript in *Observational Studies*
**Winner of the William Cochran prize for the best paper published in Observational Studies in 2015-2019*
Now: Assistant Professor, Purdue University

2015-2018 David Hirshberg, PhD Candidate in Statistics, Columbia University
Advisor
Published manuscript in *Epidemiology*
Two manuscripts in preparation
Now: Postdoctoral Fellow, Stanford University

2015-2020 Yixin Wang, PhD Candidate in Statistics, Columbia University

Published manuscript in *Biometrika*
One manuscript in preparation
**Winner of the 2018 Student Paper Award of the Biometrics Section of the American Statistical Association*
Now: Assistant Professor, University of Michigan

2015-2020 Magdalena Bennett, PhD Candidate in Education, Columbia University
Published manuscript in the *Journal of Computation and Graphical Statistics*
**Winner of the 2019 Student Paper Award of the International Biometric Society ENAR*
Now: Assistant Professor, University of Texas at Austin

Formal Teaching of Peers:
No presentations below were sponsored by outside entities

Local Invited Presentations:
No presentations below were sponsored by outside entities

2013 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Sociology Department, University of Pennsylvania

2013 New Statistical Methods for Causal Inference in Observational Studies with Applications to the Social Sciences in Health Policy
Columbia Business School, Columbia University

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Teachers College, Columbia University

2014 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, Columbia University

2016 Measuring the Effect of the Experience of Incarceration on Reoffending
Social Enterprise Leadership Forum, Columbia Business School, Columbia University

2017 New Matching Methods for Causal Inference and Impact Evaluation using Mathematical Programming
Columbia Business School

2017 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Department of Statistics, Harvard University

2017 Methods for Causal Inference to Advance Research in Health Care and Public Policy
Department of Statistics, Harvard University

2017 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Causal Inference Group, Harvard School of Public Health

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
HIV Working Group, Harvard School of Public Health

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Institute for Quantitative Social Science, Harvard University

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Harvard Catalyst Biostatistics Symposium, Harvard University

2018 New Matching Methods for Causal Inference
Harvard Business School

2018 Methods for Causal Inference to Advance Research in Health Care and Public Policy
Department of Statistics, Harvard University

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Health Economics Seminar, Harvard University

2019 What is the Impact of an Earthquake on Educational Attainment? A Matching Approach
Radcliffe Institute, Harvard University

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Kolokotrones Circle, Harvard School of Public Health

2019 Methods for Causal Inference to Advance Research in Health Care and Public Policy
Department of Statistics, Harvard University

2019 Measuring Quality of Oncology Practices
Healthcare Markets and Regulation Lab Seminar Series, Harvard Medical School

2020 Weighting for Causal Inference
Kolokotrones Symposium, Harvard School of Public Health

2020 Methods for Causal Inference to Advance Research in Health Care and Public Policy
Department of Biostatistics, Harvard School of Public Health

Report of Regional, National and International Invited Teaching and Presentations

Regional:

2012 Matching Methods in Observational Studies
Statistics Department, Columbia University

2012 Effect of the 2010 Chilean Earthquake on Posttraumatic Stress: Illustrating New Matching Methods for Clinical, Epidemiological and Health Outcomes Research
Department of Psychiatry, Columbia University

2012 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Biostatistics, Columbia University

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, Harvard University

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Institution for Social and Policy Studies, Yale University

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Robert H. Smith School of Business, University of Maryland

2017 New Matching Methods for Causal Inference and Impact Evaluation using Mathematical Programming
Operations Research Center, Massachusetts Institute of Technology

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Center for Statistical Sciences and Department of Biostatistics, Brown University

National:

2011 Contrasting Evidence Within and Between Institutions that Supply Treatment in an Observational Study of Alternative Forms of Anesthesia/Invited Presentation
Joint Statistical Meetings, Miami FL

2012 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Atlantic Causal Inference Conference, Johns Hopkins University

2012 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Joint Statistical Meetings, San Diego, CA

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, Stanford University

2013 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Biostatistics, Johns Hopkins University

2013 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Booth School of Business, University of Chicago

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Heinz College, Carnegie Mellon University

2014 Design and Analysis of Observational Studies
Kellogg School of Management, Northwestern University

2014 Optimal Matching with Direct Covariate Balance Using Integer Programming
Kaiser Permanente, San Francisco

2014 Stable Weight Adjustment for Causal Inference and Estimation with Incomplete Data
Eastern and North American Region/International Biometric Society Spring Meeting, Baltimore, MD

2014 Stable Weight Adjustment for Causal Inference and Estimation with Incomplete Data
Joint Statistical Meetings, Boston, MA

2014 Stronger Instrumental Variables Via Integer Programming for Healthcare Research
INFORMS, San Francisco, CA

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Heinz College, Carnegie Mellon University

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Biostatistics Department, Johns Hopkins University

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Statistics Department, Duke University

2015 Optimal Multilevel Matching in Clustered Observational Studies: A Case Study of the School Voucher System in Chile
SREE, Washington, DC

2015 Covariate Balanced Restricted Randomization: Optimal Designs, Exact Tests, and Asymptotic Results
INFORMS, San Francisco, CA

2016 New Matching Methods for Causal Inference and Impact Evaluation using Mathematical Programming
MEDS Kellogg School of Management, Northwestern University, Evanston, IL

2016 designmatch: Construction of Matched Samples for Randomized Experiments and Observational Studies that are Balanced by Design
Uber, California

2016 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Methods Workshop, University of California at Berkeley

2016 Optimal Designs for Causal Inference Using Integer Programming *(short course)*
Center for Mathematical Studies, Northwestern University, Evanston, IL

2016 designmatch: Construction of Matched Samples for Randomized Experiments and Observational Studies that are Balanced by Design
Atlantic Causal Inference Conference, New York, NY

2016 Maximizing the Information Content of a Balanced Matched Sample
Joint Statistical Meetings, Chicago, IL

2016 Large-scale Optimal Matching for Design-based Inference Using Integer Programming
INFORMS, Nashville, TN

2017 New Matching Methods for Causal Inference and Impact Evaluation using Mathematical Programming
Booth School of Business, University of Chicago

2017 New Matching Methods for Causal Inference and Impact Evaluation using Mathematical Programming
Department of Health Care Policy, Harvard Medical School
**Presentation was before appointment at HMS*

2017 New Matching Methods for Causal Inference and Impact Evaluation Using Mathematical Programming/Marketplace Optimization Data Science Symposium
Uber, San Francisco

2017 New Matching Methods for Causal Inference *(short course)*
Atlantic Causal Inference Conference, University of North Caroline at Chapel Hill, NC

2017 New Matching Methods for Causal Inference *(short course)*
MEDpiNet PPP Annual Meeting, Food and Drug Administration (FDA), Silver Spring, MD

2018 New Matching Methods for Causal Inference *(short course)*
Society for Research in Educational Effectiveness (SREE) Spring Conference, Washington, DC

2018 Minimal Weights for the Design of Observational Studies and Sample Surveys
Department of Biostatistics, Vanderbilt University, Nashville, TN

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Heinz College, Carnegie Mellon University, Pittsburgh, PA

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Division of Biostatistics, College of Public Health, University of Ohio, Columbus, OH

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Division of Biostatistics, College of Public Health, University of Ohio, Columbus, OH

2019 Design of Matched Studies with Improved Internal and External Validity *(short course)*
ENAR Spring Meeting, Philadelphia, PA

2019 Introduction to Causal Inference *(short course)*
Harvard Data Science Initiative Annual Conference, Cambridge, MA

2020 Multivariate Matching Methods for Generalization and Individualization
Department of Statistics, University of California, Berkeley, CA

International:

2007 A First Household Panel Survey in Chile: Methodological Considerations
Institute of Social and Economic Research, University of Essex, England

2008 How Income Stratification is Perpetuated Across Generations? The Contribution of Longitudinal Surveys
Expansiva Workshop, Santiago, Chile

2011 Contrasting Evidence Within and Between Institutions that Supply Treatment in an Observational Study of Alternative Forms of Anesthesia
International Conference on Health Policy Statistics, Cleveland, OH

2012 Estimation of the Effect of Prophylactic CPAP on Very Low Birth Weight Infants Using Matching
Neocosur Conference, Buenos Aires, Argentina

2012 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, Warwick University, England

2012 Designing an Observational Study to be Less Sensitive to Unmeasured Biases: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Biostatistics and Epidemiology, McGill University

2013 Matching for Balance, Pairing for Heterogeneity in an Observational Study of Effectiveness of For-profit and Not-for-profit High Schools in Chile
Joint Statistical Meetings, Montreal, CA

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, University of Oxford, England

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Department of Statistics, London School of Economics, England

2013 Using Mixed Integer Programming for Matching in Observational Studies: Effect of the 2010 Chilean Earthquake on Posttraumatic Stress
Statistical Laboratory, University of Cambridge, England

2013 Estimation Strategies in Observational Studies
Neocosur Conference, Buenos Aires, Argentina

2013 Effect of Prophylactic CPAP in Very Low Birth Weight Infants in South America
Neocosur Conference, Buenos Aires, Argentina

2014 Instrumental Variable for Causal Inference in the Health Sciences
Neocosur Conference, Buenos Aires, Argentina

2014 New Methods for Causal Inferences in the Health and Social Sciences *(short course)*
Columbia Global Center/Universidad Catolica de Chile, Santiago, Chile

2015 Design of Observational Studies *(short course)*
International Workshop on Applied Statistics, Bogota, Columbia

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
Statistical Laboratory, University of Cambridge, England

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
UK Causal Inference Meeting, University of Bristol, England

2015 Stable Weights that Balance Covariates for Causal Inference and Estimation with Incomplete Data
International Workshop on Applied Statistics, Bogota, Colombia

2016 Causal Inference and Impact Evaluation in Observational Studies: New Matching Methods to Approximate a Randomized Experiment
School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

2016 Toward an Evaluation of the Comparative Effectiveness of the Intensive Care Units of the Neocosur Network
Neocosur Conference, Buenos Aires, Argentina

2017 New Matching Methods for Causal Inference
United Kingdom Causal Inference Meeting, University of Essex, England

2017 Comparative Effectiveness of the Intensive Care Units of the Neocosur Network through Weighted Samples
Neocosur Conference, Buenos Aires, Argentina

2017 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
International Conference of the ERCIM WG on Computational and Methodological Statistics, Senate House, University of London, England

2017 New Statistical Methods for Causal Inference in Medicine and Public Health
School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

2017 New Statistical Methods for Causal Inference in Medicine and Public Health
Ministry of Health, Gobierno de Chile, Santiago, Chile

2018 Minimal Approximate Balancing Weights: Asymptotic Properties and Practical Considerations
International Conference in Health Policy Statistics, Charleston, SC

2018 Recent Developments in Causal Inference
International Conference in Health Policy Statistics, Charleston, SC

2018 New Matching Methods to Increase the Internal and External Validity of Observational Studies
WHO World Mental Health Surveys Annual Meeting, Boston, MA

2018 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Joint Statistical Meetings, Vancouver, Canada

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
School of Medicine, University of Nottingham, England

2019 General Discontinuity Designs Using Covariates
Statistical Laboratory, University of Cambridge, England

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, England

2019 Introduction to Causal Inference
Harvard Data Science Annual Conference, Boston, MA

2019 Building Representative Matched Samples in Large-Scale Observational Studies with Multivalued Treatments
Optimization-Conscious Econometrics Conference, University of Chicago, Chicago, IL

2019 Complex Discontinuity Designs Using Covariates
The Statistical and Applied Mathematical Sciences Institute (SAMSI) Opening Workshop on Causal Inference, Durham, NC

2019 Complex Discontinuity Designs Using Covariates
International Conference of the ERCIM WG on Computational and Methodological Statistics, Senate House, University of London, England

2020 Matching Techniques for Generalization and Individualization
Joint Statistical Meetings, Philadelphia, PA

2020 Effectiveness of Localized Lockdowns in the COVID-19 Pandemic
COVID-19 Modeling Workshop, Santiago, Chile

Report of Scholarship

Peer-Reviewed Scholarship in print or other media:

Research Investigations:

1) **Zubizarreta JR**, Reinke CE, Kelz RR, Silber JH, Rosenbaum PR. Matching for Several Sparse Nominal Variables in a Case Control Study of Readmission Following Surgery. The American Statistician 2011. 65: 229-238.

2) **Zubizarreta JR**, Neuman MD, Silber JH, Rosenbaum PR. Contrasting Evidence Within and between Institutions that Supply Treatment in an Observational Study of Alternative Forms of Anesthesia. Journal of the American Statistical Association 2012. 107: 901-915.

3) **Zubizarreta JR.** Using Mixed Integer Programming for Matching in an Observational Study of Acute Kidney Injury after Surgery. Journal of the American Statistical Association 2012. 107: 1360-1371.

4) Reinke CE, Kelz RR, **Zubizarreta JR**, Lanyu M, Saynisch P, Kyle FA, Even-Shoshan O, Fleisher LA, Silber JH. Obesity and Readmission in Elderly Surgical Patients. Surgery 2012. 152: 355-362.

5) **Zubizarreta JR**, Small DS, Goyal NK, Lorch SA, Rosenbaum PR. Stronger Instruments Via Integer Programming in an Observational Study of Late Preterm Birth Outcome. Annals of Applied Statistics 2013. 7: 25-50.

6) **Zubizarreta JR**, Cerda M, Rosenbaum PR. Effect of the 2010 Chilean Earthquake on Posttraumatic Stress: Reducing Sensitivity to Unmeasured Bias Through Study Design. Epidemiology 2013. 24: 79-87 (with discussion). Winner of the Rothman Prize for the best paper published in Epidemiology in 2013.

7) Goyal NK, **Zubizarreta JR**, Small DS, Lorch SA. Length of Stay and Readmission Risk for late Preterm Infants: An Instrumental Variable Approach. Hospital Pediatrics 2013. 3: 7-15.

8) Kelz RR, Reinke CE, **Zubizarreta JR**, Wang M, Saynisch P, Reese P, Even-Shoshan O, Reese PR, Fleisher LA, Silber JH, Acute Kidney Injury, Renal Function, and the Elderly Obese Surgical Patient: A Matched Case-Control Study. Annals of Surgery 2013. 258: 359-363.

9) Yang F, **Zubizarreta JR**, Small DS, Lorch SA, Rosenbaum PR. Dissonant Conclusions When Testing the Validity of an Instrumental Variable. The American Statistician 2014. 68: 253-263.

10) **Zubizarreta JR**, Paredes RD, Rosenbaum PR. Matching for Balance, Pairing for Heterogeneity in an Observational Study of the Effectiveness of For-Profit and Not-for-profit High Schools in Chile. Annals of Applied Statistics 2014. 8: 2096-2121.

11) **Zubizarreta JR**, Small DS, Rosenbaum PR. Isolation in the Construction of Natural Experiments. Annals of Applied Statistics 2014. 8: 2096-2121.

12) Neuman MD, Rosenbaum PR, Ludwig JM, **Zubizarreta JR**, Silber JH. Anesthesia Technique, Mortality, and Length of Stay After Hip Fracture Surgery. Journal of the American Medical Association 2014. 311: 2508-2517.

13) Keele L, Titiunik R, **Zubizarreta JR.** Enhancing a Geographic Regression Discontinuity Design Through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout. Journal of the Royal Statistical Society: Series A 2015. 178: 223-239.

14) Hsu J, **Zubizarreta JR**, Small DS, Rosenbaum PR. Strong Control of the Family-Wise Error Rate in Observational Studies that Discover Effect Modification by Exploratory Methods. Biometrika 2015. 102: 767-782.

15) **Zubizarreta JR.** Stable Weights that Balance Covariates for Estimation with Incomplete Outcome Data. Journal of the American Statistical Association 2015. 110: 910-922.

16) Resa MA, **Zubizarreta JR.** Evaluation of Subset Matching Methods and Forms of Covariate Balance. Statistics in Medicine 2016. 35: 4961-4979.

17) Kilcioglu C, **Zubizarreta JR.** Maximizing the Information Content of a Balanced Matched Sample in a Study of the Economic Performance of Green Buildings. Annals of Applied Statistics 2016. 10: 1997-2020.

18) **Zubizarreta JR**, Lorch SA, Marshall G, D'Apremont I, Tapia JL. Effect of Prophylactic CPAP in Very Low Birth Weight Infants in South America. Journal of Perinatology 2016. 36: 629-634.

19) **Zubizarreta JR**, Keele L. Optimal Multilevel Matching in Clustered Observational Studies: A Case Study of the School Voucher System in Chile. Journal of the American Statistical Association 2017. 112: 547-560.

20) Hirshberg, DA, **Zubizarreta JR**. On Two Approaches to Weighting in Causal Inference. Invited commentary. Epidemiology 2017. 28: 812-816.

21) Rosellini AJ, Dussaillant F, **Zubizarreta JR**, Kessler R, Rose S. Predicting Post-traumatic Stress Disorder Following a Natural Disaster. Journal of Psychiatric Research 2018. 96: 15-22.

22) Hu W, Chan C, **Zubizarreta JR**, Escobar G. An Examination of Early Transfers to the ICU Based on a Physiologic Risk Score. Manufacturing & Service Operations Management 2018. 20.3: 531-549.

23) Hu W, Chan C, **Zubizarreta JR**, Escobar G. Incorporating Longitudinal Comorbidity and Acute Physiology Data in Template Matching for Assessing Hospital Quality: An Exploratory Study in an Integrated Health Care Delivery System. Medical Care 2018. 56: 448-454

24) Haneuse S, **Zubizarreta JR**, Normand SLT. On Assessing the Quality of Health Care Providers Using Time-Varying Patient Outcomes. Invited commentary. Biometrics 2018. 74: 1395-1397.

25) **Zubizarreta JR**, Small DS, Rosenbaum PR. A Simple Example of Isolation in Building a Natural Experiment. Chance 2018. 31: 16-23.

26) Aronow PM, Crawford FW, **Zubizarreta JR**. Confidence Intervals for Linear Unbiased Estimators Under Constrained Dependence. Electronic Journal of Statistics 2018. 12: 2238-2252.

27) Visconti G, **Zubizarreta JR**. Handling Limited Overlap in Observational Studies with Cardinality Matching. Observational Studies 2018. 4:217-249.

28) Miratrix LW, Wager S, **Zubizarreta JR**. Shape-Constrained Partial Identification of a Population Mean Under Unknown Probabilities of Sample Selection. Biometrika 2018. 107: 103-114.

29) Gondi S, Wright AA, Landrum MB, **Zubizarreta JR**, Chernew ME, Keating N. Multimodality Cancer Care and Implications for Episode Payments in Cancer. American Journal of Managed Care 2019; 25: 294-295.

30) Kessler RC, Bossarte RM, Luedtke A, Zaslavsky AM, **Zubizarreta JR**, Machine Learning Methods for Developing Precision Treatment Rules, Behaviour Research and Therapy 2019. 120, 103412.

31) Kessler RC, Bossarte RM, Luedtke A, Zaslavsky AM, **Zubizarreta JR**, Suicide Prediction Models: A Critical Review of Recent Research with Recommendations for the Way Forward, Molecular Psychiatry 2019. 1-12.

32) Wang Y, **Zubizarreta JR**, Minimal Dispersion Approximately Balancing Weights: Asymptotic Properties and Practical Considerations. Biometrika, 2020. 107: 93-105.

33) Wilcock AD, Schwamm LH, Zachrison KS, Uscher-Pines L, **Zubizarreta JR**, Mehrotra A. Trends in Care Delivery and Outcomes for Acute Stroke and Transient Ischemic Attacks among Rural and Urban Medicare Beneficiaries, 2008-2017, Journal of the American Medical Association, Neurology, 2020; 77: 863-871.

34) Resa MA, **Zubizarreta JR**. Direct and Stable Weight Adjustment in Non-Experimental Studies with Multi-valued Treatments: Analysis of the Impact of an Earthquake on Posttraumatic Stress. Journal of the Royal Statistical Society: Series A, 2020; 183: 1387-1410.

35) Bennett M, Vielma, JP, **Zubizarreta JR**. Building a Representative Matched Samples with Multi-valued Treatments in Large Observational Studies. Journal of Computational and Graphical Statistics, in press.

36) Chattopadhyay A, Hase C, **Zubizarreta JR**. Balancing Versus Modeling Approaches to Weighting in Practice. Statistics in Medicine, 2020; 39: 3227-3254.

Non-peer reviewed scholarship in print or other media:
Proceedings of meetings or other non-peer reviewed scholarship:

1) **Zubizarreta JR.** Educational Outcomes of the Children of the Poor: the Chilean Case, background paper prepared for the Human Development Report 2008/2009 for Latin America and the Caribbean. United Nations Development Plan 2009.

2) **Zubizarreta JR.** Introduction to matching in Observational Studies Using Mixed Integer Programming. Transactions of the Annual Deming Conference on Applied Statistics 2012.

Reviews, chapters, monographs and editorials:

1) Rodriguez C, Dominguez P, Undurraga E, **Zubizarreta JR**. Identification and Characterization of Vulnerable Populations: Elements for the Measurement of Risk. In: Towards the Bicentenary: Proposals for Chile. Ediciones UC 2009. 305-328. (In Spanish.)

2) Paredes RD, Volante P, Opazo MO, **Zubizarreta JR**. Shared Financing in the Chilean Subsidized Education. In: Proposals for Chile. Ediciones UC 2013. 51-83. (In Spanish.)

3) Mozer R, Rubin DB, **Zubizarreta, JR**. Statistical Inference for Causal Effects in Clinical Psychology: Fundamental Concepts and Analytical Approaches. In: Handbook of Research Methods in Clinical Psychology, in press.

Thesis:

Zubizarreta JR. Optimal Designs for Observational Studies Using Integer Programming. University of Pennsylvania. 2013.

Report of Technological and Other Scientific Innovations

Statistical Software:

1) depinf package for R, with Peter Aronow (Yale) and Forrest Crawford (Yale): statistical package for the construction of confidence intervals for linear unbiased estimators under constrained dependence.
2) designmatch package for R, with Cinar Kilcioglu (Uber) and Juan Pablo Vielma (MIT): statistical package for the construction of matched samples that are balanced and representative by design. *21534 downloads as of July, 2020.*
3) mipmatch package for R: statistical package for the construction of matched samples using mixed integer programming.
4) sbw package for R, with Yige Li (HCP) and Mohammed-Amine Allouah (Columbia): statistical package for the construction of stable weights that balance covariates for causal inference and estimation with incomple outcome data. *3149 downloads as of July, 2020.*
5) scbounds package for R, with Luke Miratrix (HGSE) and Stefan Wager (Stanford): statistical package for the construction of shape-constrained bounds for a population mean under unknown probabilities of sample selection.