Special Issue on Causal Inference

Editorial: Special Issue on “Causal Inference” ............... Jason Roy and Dylan Small 335
Matching Methods for Observational Studies Derived from Large Administrative Databases ..................... Ruqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum 338
Comment: Matching Methods for Observational Studies Derived from Large Administrative Databases ............. Fredrik Sävje 356
Comment: Matching Methods for Observational Studies Derived from Large Administrative Databases ... Mark M. Fredrickson, Josh Errickson and Ben B. Hansen 361
Commentary on Yu et al.: Opportunities and Challenges for Matching Methods in Large Databases .................... Elizabeth A. Stuart and Benjamin Ackerman 367
Rejoinder: Matching Methods for Observational Studies Derived from Large Administrative Databases .... Ruqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum 371
Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study .......................... Tianchen Qian, Predrag Klasnja and Susan A. Murphy 375
Comment: Clarifying Endogeneous Data Structures and Consequent Modelling Choices Using Causal Graphs ............ Erica E. M. Moodie and David A. Stephens 391
Moving Toward Rigorous Evaluation of Mobile Health Interventions ........ Kristin A. Linn 394
Rejoinder: Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study .......................... Tianchen Qian, Predrag Klasnja and Susan A. Murphy 400
Invariance, Causality and Robustness ............................... Peter Bühlmann 404
Comment: Invariance, Causality and Robustness by P. Bühlmann .......... Vanessa Didelez 427
Comment: Invariance and Causal Inference ............................ Stefan Wager 430
Rejoinder: Invariance, Causality and Robustness ............................... Peter Bühlmann 434
Outcome-Wide Longitudinal Designs for Causal Inference: A New Template for Empirical Studies .................... Tyler J. VanderWeele, Maya B. Mathur and Ying Chen 437
Comment: On the Potential for Misuse of Outcome-Wide Study Designs, and Ways to Prevent It .......................... Stijn Vansteelandt and Oliver Dukes 467
Comment: Outcome-Wide Individualized Treatment Strategies .................................................. Ashkan Ertefaie and Brent A. Johnson 472

Statistical Science [ISSN 0883-4237 (print); ISSN 2168-8745 (online)], Volume 35, Number 3, August 2020. Published quarterly by the Institute of Mathematical Statistics, 3163 Somerset Drive, Cleveland, OH 44122, USA. Periodicals postage paid at Cleveland, Ohio and at additional mailing offices.

POSTMASTER: Send address changes to Statistical Science, Institute of Mathematical Statistics, Dues and Subscriptions Office, 9650 Rockville Pike—Suite L2310, Bethesda, MD 20814-3998, USA.

Copyright © 2020 by the Institute of Mathematical Statistics
Printed in the United States of America
A New Template for Empirical Studies: From positivity to Positivity ..........Rhian Daniel 476
Rejoinder: The Future of Outcome-Wide Studies ..................................................Tyler J. VanderWeele, Maya B. Mathur and Ying Chen 479
A Nonparametric Super-Efficient Estimator of the Average Treatment Effect 484
David Benkeser, Weixin Cai and Mark J. van der Laan 484
Comment: Increasing Real World Usage of Targeted Minimum Loss-Based Estimators ..........................................................Mireille E. Schnitzer 496
Comment: Automated Analyses: Because We Can, Does It Mean We Should? ..........................................................Susan M. Shortreed and Erica E. M. Moodie 499
Comment: Stabilizing the Doubly-Robust Estimators of the Average Treatment Effect under Positivity Violations .................................................Fan Li 503
Rejoinder: A Nonparametric Superefficient Estimator of the Average Treatment Effect ..................................................David Benkeser, Weixin Cai and Mark J. van der Laan 511
On Nearly Assumption-Free Tests of Nominal Confidence Interval Coverage for Causal Parameters Estimated by Machine Learning
Lin Liu, Rajarshi Mukherjee and James M. Robins 518
Discussion of “On Nearly Assumption-Free Tests of Nominal Confidence Interval Coverage for Causal Parameters Estimated by Machine Learning” Edward H. Kennedy, Sivaraman Balakrishnan and Larry Wasserman 540
Rejoinder: On nearly assumption-free tests of nominal confidence interval coverage for causal parameters estimated by machine learning
Lin Liu, Rajarshi Mukherjee and James M. Robins 545
Special Issue on Causal Inference

Editorial: Special Issue on “Causal Inference”
   Jason Roy and Dylan Small

Matching Methods for Observational Studies Derived from Large Administrative Databases
   Ruoqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum

Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study
   Tianchen Qian, Predrag Klasnja and Susan A. Murphy

Invariance, Causality and Robustness
   Peter Bühlmann

Outcome-Wide Longitudinal Designs for Causal Inference: A New Template for Empirical Studies
   Tyler J. VanderWeele, Maya B. Mathur and Ying Chen

A Nonparametric Super-Efficient Estimator of the Average Treatment Effect
   David Benkeser, Weixin Cai and Mark J. van der Laan

On Nearly Assumption-Free Tests of Nominal Confidence Interval Coverage for Causal Parameters Estimated by Machine Learning
   Lin Liu, Rajarshi Mukherjee and James M. Robins
Editorial: Special Issue on “Causal Inference”

Jason Roy and Dylan Small

REFERENCES

Statist. Assoc. 81 945–970. MR0867618

ROSENBAUM, P. R. and RUBIN, D. B. (1983). The central role of
the propensity score in observational studies for causal effects.
Biometrika 70 41–55. MR0742974 https://doi.org/10.1093/biomet/
70.1.41
Matching Methods for Observational Studies Derived from Large Administrative Databases

Ruoqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum

Abstract. We propose new optimal matching techniques for large administrative data sets. In current practice, very large matched samples are constructed by subdividing the population and solving a series of smaller problems, for instance, matching men to men and separately matching women to women. Without simplification of some kind, the time required to optimally match $T$ treated individuals to $T$ controls selected from $C \geq T$ potential controls grows much faster than linearly with the number of people to be matched—the required time is of order $O((T + C)^3)$—so splitting one large problem into many small problems greatly accelerates the computations. This common practice has several disadvantages that we describe. In its place, we propose a single match, using everyone, that accelerates the computations in a different way. In particular, we use an iterative form of Glover’s algorithm for a doubly convex bipartite graph to determine an optimal caliper for the propensity score, radically reducing the number of candidate matches; then we optimally match in a large but much sparser graph. In this graph, a modified form of near-fine balance can be used on a much larger scale, improving its effectiveness. We illustrate the method using data from US Medicaid, matching children receiving surgery at a children’s hospital to similar children receiving surgery at a hospital that mostly treats adults. In the example, we form 38,841 matched pairs from 159,527 potential controls, controlling for 29 covariates plus 463 Principal Surgical Procedures, plus 973 Principal Diagnoses. The method is implemented in an R package bigmatch available from CRAN.

Key words and phrases: Causal inference, fine balance, Glover’s algorithm, observational study, optimal caliper, optimal matching, propensity score.

REFERENCES


Ruoqi Yu is a PhD student, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6340, USA (e-mail: ruoqiyu@wharton.upenn.edu). Jeffrey H. Silber is Professor, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA (e-mail: silber@email.chop.edu; URL: https://cor.research.chop.edu/). Paul R. Rosenbaum is Professor, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6340, USA (e-mail: rosenbaum@wharton.upenn.edu).


Comment: Matching Methods for Observational Studies Derived from Large Administrative Databases

Fredrik Sävje

REFERENCES


Fredrik Sävje is Assistant Professor, Department of Political Science and Department of Statistics and Data Science, Yale University, Rosenkranz Hall, 115 Prospect Street, New Haven, Connecticut 06520, USA (e-mail: fredrik.savje@yale.edu).
Comment: Matching Methods for Observational Studies Derived from Large Administrative Databases

Mark M. Fredrickson, Josh Errickson and Ben B. Hansen

REFERENCES


Commentary on Yu et al.: Opportunities and Challenges for Matching Methods in Large Databases

Elizabeth A. Stuart and Benjamin Ackerman

REFERENCES


Elizabeth A. Stuart is Professor of Mental Health, Biostatistics, and Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, 21205, USA (e-mail: estuart@jhu.edu). Benjamin Ackerman is a Doctoral Candidate in the Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, 21205, USA.
Rejoinder: Matching Methods for Observational Studies Derived from Large Administrative Databases

Ruoqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum

REFERENCES


Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study

Tianchen Qian, Predrag Klasnja and Susan A. Murphy

Abstract. Mobile health is a rapidly developing field in which behavioral treatments are delivered to individuals via wearables or smartphones to facilitate health-related behavior change. Micro-randomized trials (MRT) are an experimental design for developing mobile health interventions. In an MRT, the treatments are randomized numerous times for each individual over course of the trial. Along with assessing treatment effects, behavioral scientists aim to understand between-person heterogeneity in the treatment effect. A natural approach is the familiar linear mixed model. However, directly applying linear mixed models is problematic because potential moderators of the treatment effect are frequently endogenous—that is, may depend on prior treatment. We discuss model interpretation and biases that arise in the absence of additional assumptions when endogenous covariates are included in a linear mixed model. In particular, when there are endogenous covariates, the coefficients no longer have the customary marginal interpretation. However, these coefficients still have a conditional-on-the-random-effect interpretation. We provide an additional assumption that, if true, allows scientists to use standard software to fit linear mixed model with endogenous covariates, and person-specific predictions of effects can be provided. As an illustration, we assess the effect of activity suggestion in the HeartSteps MRT and analyze the between-person treatment effect heterogeneity.

Key words and phrases: Linear mixed model, endogenous covariates, micro-randomized trial, causal inference.

REFERENCES


Tianchen Qian is Postdoctoral Fellow, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: qiantianchen@fas.harvard.edu). Predrag Klasnja is Assistant Professor, School of Information, University of Michigan, Ann Arbor, Massachusetts 48109, USA (e-mail: klasnja@umich.edu). Susan A. Murphy is Professor, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: samurphy@fas.harvard.edu).
Comment: Clarifying Endogeneous Data Structures and Consequent Modelling Choices Using Causal Graphs

Erica E. M. Moodie and David A. Stephens

REFERENCES


Moving Toward Rigorous Evaluation of Mobile Health Interventions

Kristin A. Linn

Abstract. Qian, Klasnja and Murphy provide an assumption that allows for unbiased estimation of treatment effects in microrandomized trials when the data are modeled using linear mixed models with endogenous covariates. In this discussion, the validity of the assumption in the context of the HeartSteps microrandomized trial is reassessed. The utility of a marginal interpretation, versus the proposed conditional-on-the-random-effect interpretation, is also discussed.

Key words and phrases: Linear mixed model, endogenous covariates, microrandomized trial, causal inference.

REFERENCES


Kristin A. Linn is Assistant Professor of Biostatistics, Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, 423 Guardian Drive, Philadelphia, Pennsylvania 19104, USA (e-mail: klinn@upenn.edu).
Comment: Diagnostics and Kernel-based Extensions for Linear Mixed Effects Models with Endogenous Covariates

Hunyong Cho, Joshua P. Zitovsky, Xinyi Li, Minxin Lu, Kushal Shah, John Sperger, Matthew C. B. Tsilimigras and Michael R. Kosorok

Abstract. We discuss “Linear mixed models with endogenous covariates: modeling sequential treatment effects with application to a mobile health study” by Qian, Klasnja and Murphy. In this discussion, we study when the linear mixed effects models with endogenous covariates are feasible to use by providing examples and diagnostic tools as well as discussing potential extensions. This includes evaluating feasibility of partial likelihood-based inference, checking the conditional independence assumption, estimation of marginal effects, and kernel extensions of the model.

Key words and phrases: Linear mixed models, partial likelihood, conditional independence test, marginal effects, kernel mixed models.

REFERENCES


Hunyong Cho is a graduate student, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: hunycho@live.unc.edu). Joshua P. Zitovsky is a graduate student, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: joshz@live.unc.edu). Xinyi Li is a postdoctoral fellow, Statistical and Applied Mathematical Sciences Institute (SAMSI) and University of North Carolina at Chapel Hill, Durham/Chapel Hill, North Carolina 27516, USA (e-mail: xli@samsi.info). Minxin Lu is a graduate student, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: mino12@live.unc.edu). Kushal Shah is a graduate student, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: kushshah@live.unc.edu). John Sperger is a graduate student, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: jsperger@live.unc.edu). Matthew C. B. Tsilimigras is a postdoctoral fellow, Department of Epidemiology, Department of Nutrition, Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: matthew_tsilimigras@unc.edu). Michael R. Kosorok is the W.R. Kenan, Jr. Distinguished Professor and Chair, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, USA (e-mail: kosorok@unc.edu).
Rejoinder: Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study

Tianchen Qian, Predrag Klasnja and Susan A. Murphy

REFERENCES


Invariance, Causality and Robustness
2018 Neyman Lecture

Peter Bühlmann

Abstract. We discuss recent work for causal inference and predictive robustness in a unifying way. The key idea relies on a notion of probabilistic invariance or stability: it opens up new insights for formulating causality as a certain risk minimization problem with a corresponding notion of robustness. The invariance itself can be estimated from general heterogeneous or perturbation data which frequently occur with nowadays data collection. The novel methodology is potentially useful in many applications, offering more robustness and better “causal-oriented” interpretation than machine learning or estimation in standard regression or classification frameworks.

Key words and phrases: Anchor regression, causal regularization, distributional robustness, heterogeneous data, instrumental variables regression, interventional data, Random Forests, variable importance.

REFERENCES


Comment: Invariance, Causality and Robustness

Vanessa Didelez

REFERENCES


Vanessa Didelez is Professor, Leibniz Institute for Prevention Research and Epidemiology—BIPS, and Faculty of Mathematics and Computer Science, University of Bremen, Germany (e-mail: didelez@leibniz-bips.de).
Comment: Invariance and Causal Inference

Stefan Wager

REFERENCES


*Econometrica* **86** 591–616. MR3783340 https://doi.org/10.3982/ECTA13288


Stefan Wager is Assistant Professor of Operations, Information and Technology, and Assistant Professor of Statistics (by courtesy), Graduate School of Business, Stanford University, Stanford, California 94305, USA (e-mail: swager@stanford.edu).
Rejoinder: Invariance, Causality and Robustness

Peter Bühlmann

Abstract. We sincerely thank Vanessa Didelez and Stefan Wager for their insightful and inspiring comments. Their views and thoughts on the topic of my article are of great value and truly contribute to put it into greater perspective.

Key words and phrases: Anchor regression, causal regularization, distributional robustness, heterogeneous treatment effects, instrumental variables regression, random forests.

REFERENCES


Outcome-Wide Longitudinal Designs for Causal Inference: A New Template for Empirical Studies

Tyler J. VanderWeele, Maya B. Mathur and Ying Chen

Abstract. In this paper, we propose a new template for empirical studies intended to assess causal effects: the outcome-wide longitudinal design. The approach is an extension of what is often done to assess the causal effects of a treatment or exposure using confounding control, but now, over numerous outcomes. We discuss the temporal and confounding control principles for such outcome-wide studies, metrics to evaluate robustness or sensitivity to potential unmeasured confounding for each outcome and approaches to handle multiple testing. We argue that the outcome-wide longitudinal design has numerous advantages over more traditional studies of single exposure-outcome relationships including results that are less subject to investigator bias, greater potential to report null effects, greater capacity to compare effect sizes, a tremendous gain in the efficiency for the research community, a greater policy relevance and a more rapid advancement of knowledge. We discuss both the practical and theoretical justification for the outcome-wide longitudinal design and also the pragmatic details of its implementation, providing publicly available R code.

Key words and phrases: Causal inference, confounding, multiple testing, sensitivity analysis, bias, longitudinal data.

REFERENCES


Tyler J. VanderWeele is John L. Loeb and Frances Lehman Loeb Professor of Epidemiology, Department of Epidemiology, Harvard University, Boston, Massachusetts, USA (e-mail: tvanderw@hsph.harvard.edu). Maya B. Mathur is Postdoctoral Research Fellow, Department of Epidemiology, Harvard University, Boston, Massachusetts, USA (e-mail: mmathur@stanford.edu). Ying Chen is Research Scientist, Institute for Quantitative Social Science, Harvard University, Boston, Massachusetts, USA (e-mail: yic867@mail.harvard.edu).
Comment: On the Potential for Misuse of Outcome-Wide Study Designs, and Ways to Prevent It

Stijn Vansteelandt and Oliver Dukes

REFERENCES


Dukes, O., Avagyan, V. and Vansteelandt, S. (2020). Doubly robust tests of exposure effects under high-dimensional confound-
Comment: Outcome-Wide Individualized Treatment Strategies

Ashkan Ertefaie and Brent A. Johnson

REFERENCES


Ashkan Ertefaie is an Assistant Professor, Department of Biostatistics and Computational Biology, University of Rochester, Rochester, New York 14642, USA (e-mail: ashkan_ertefaie@urmc.rochester.edu). Brent A. Johnson is a Professor, Department of Biostatistics and Computational Biology, University of Rochester, Rochester, New York 14642, USA (e-mail: brent_johnson@urmc.rochester.edu).
A New Template for Empirical Studies: From positivity to Positivity

Rhian Daniel

REFERENCES


Rejoinder: The Future of Outcome-Wide Studies

Tyler J. VanderWeele, Maya B. Mathur and Ying Chen

REFERENCES


A Nonparametric Super-Efficient Estimator of the Average Treatment Effect

David Benkeser, Weixin Cai and Mark J. van der Laan

Abstract. Doubly robust estimators are a popular means of estimating causal effects. Such estimators combine an estimate of the conditional mean of the outcome given treatment and confounders (the so-called outcome regression) with an estimate of the conditional probability of treatment given confounders (the propensity score) to generate an estimate of the effect of interest. In addition to enjoying the double-robustness property, these estimators have additional benefits. First, flexible regression tools, such as those developed in the field of machine learning, can be utilized to estimate the relevant regressions, while the estimators of the treatment effects retain desirable statistical properties. Furthermore, these estimators are often statistically efficient, achieving the lower bound on the variance of regular, asymptotically linear estimators. However, in spite of their asymptotic optimality, in problems where causal estimands are weakly identifiable, these estimators may behave erratically. We propose new estimation techniques for use in these challenging settings. Our estimators build on two existing frameworks for efficient estimation: targeted minimum loss estimation and one-step estimation. However, rather than using an estimate of the propensity score in their construction, we instead opt for an alternative regression quantity when building our estimators: the conditional probability of treatment given the conditional mean outcome. We discuss the theoretical implications and demonstrate the estimators’ performance in simulated and real data.

Key words and phrases: Causal inference, average treatment effect, asymptotic linearity, efficient influence function, collaborative targeted minimum loss estimation, super efficiency.

REFERENCES


David Benkeser is an Assistant Professor at Emory University in the Rollins School of Public Health, Department of Biostatistics and Bioinformatics, Atlanta, Georgia 30322, USA (e-mail: benkeser@emory.edu). Weixin Cai is a quantitative researcher at Citadel LLC, Seattle, Washington 98122, USA. Mark J. van der Laan is the Jiann-Ping Hsu/Karl E. Peace Professor of Biostatistics and Statistics at the University of California, Berkeley, Berkeley, California 94720, USA.
Comment: Increasing Real World Usage of Targeted Minimum Loss-Based Estimators

Mireille E. Schnitzer

REFERENCES


Mireille E. Schnitzer is Associate Professor of Biostatistics, Faculty of Pharmacy and School of Public Health, Université de Montréal, 2940, chemin de Polytechnique, Montréal, Québec, Canada, H3T 1J4 (e-mail: mireille.schnitzer@umontreal.ca).
Comment: Automated Analyses: Because We Can, Does It Mean We Should?

Susan M. Shortreed and Erica E. M. Moodie

REFERENCES


Comment: Stabilizing the Doubly-Robust Estimators of the Average Treatment Effect under Positivity Violations

Fan Li

Abstract. Doubly-robust estimators within the one-step and TMLE frameworks could exhibit finite-sample bias and excess variability under positivity violations. We comment on how the application of a stabilization factor may improve the efficiency property of one-step estimator and TMLE, and the comparisons with their collaborative counterparts using the adaptive propensity scores.

Key words and phrases: Efficient influence function, overlap weighting, trimming, one-step estimation, targeted maximum likelihood estimation.

REFERENCES


Fan Li is Assistant Professor, Department of Biostatistics, Yale School of Public Health, 135 College Street, Suite 200, New Haven, Connecticut 06510, USA (e-mail: fan.f.li@yale.edu).
Rejoinder: A Nonparametric Superefficient Estimator of the Average Treatment Effect

David Benkeser, Weixin Cai and Mark J. van der Laan

REFERENCES


HARTNETT, K. (2018). To build truly intelligent machines, teach them cause and effect [online; accessed 13-January-2020].


On Nearly Assumption-Free Tests of Nominal Confidence Interval Coverage for Causal Parameters Estimated by Machine Learning

Lin Liu, Rajarshi Mukherjee and James M. Robins

Abstract. For many causal effect parameters of interest, doubly robust machine learning (DRML) estimators $\hat{\psi}_1$ are the state-of-the-art, incorporating the good prediction performance of machine learning; the decreased bias of doubly robust estimators; and the analytic tractability and bias reduction of sample splitting with cross-fitting. Nonetheless, even in the absence of confounding by unmeasured factors, the nominal $(1 - \alpha)$ Wald confidence interval $\hat{\psi}_1 \pm z_{\alpha/2} \hat{\text{se}}[\hat{\psi}_1]$ may still undercover even in large samples, because the bias of $\hat{\psi}_1$ may be of the same or even larger order than its standard error of order $n^{-1/2}$.

In this paper, we introduce essentially assumption-free tests that (i) can falsify the null hypothesis that the bias of $\hat{\psi}_1$ is of smaller order than its standard error, (ii) can provide a upper confidence bound on the true coverage of the Wald interval, and (iii) are valid under the null under no smoothness/sparsity assumptions on the nuisance parameters. The tests, which we refer to as Assumption Free Empirical Coverage Tests (AFECTs), are based on a U-statistic that estimates part of the bias of $\hat{\psi}_1$.

Our claims need to be tempered in several important ways. First no test, including ours, of the null hypothesis that the ratio of the bias to its standard error is smaller than some threshold $\delta$ can be consistent [without additional assumptions (e.g., smoothness or sparsity) that may be incorrect]. Second, the above claims only apply to certain parameters in a particular class. For most of the others, our results are unavoidably less sharp. In particular, for these parameters, we cannot directly test whether the nominal Wald interval $\hat{\psi}_1 \pm z_{\alpha/2} \hat{\text{se}}[\hat{\psi}_1]$ undercovers. However, we can often test the validity of the smoothness and/or sparsity assumptions used by an analyst to justify a claim that the reported Wald interval’s actual coverage is no less than nominal. Third, in the main text, with the exception of the simulation study in Section 1, we assume we are in the semisupervised data setting (wherein there is a much larger dataset with information only on the covariates), allowing us to regard the covariance matrix of the covariates as known. In the simulation in Section 1, we consider the setting in which estimation of the covariance matrix is required. In the simulation, we used a data adaptive estimator which performs very well in our simulations, but the estimator’s theoretical sampling behavior remains unknown.

Key words and phrases: Causal inference, assumption-free, valid inference, U-statistics, higher-order influence functions.

Lin Liu is Assistant Professor, Institute of Natural Sciences, School of Mathematical Sciences and SJTU-Yale Joint Center for Biostatistics, Shanghai Jiao Tong University, Shanghai 200420, China (e-mail: linliu@alumni.tongji.edu.cn). Rajarshi Mukherjee is Assistant Professor, Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts 02115, USA (e-mail: ram521@mail.harvard.edu). James M. Robins is Mitchell L. and Robin LaFoley Dong Professor, Department of
REFERENCES


Epidemiology and Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts 02115, USA (e-mail: robins@hsph.harvard.edu).
Discussion of “On Nearly Assumption-Free Tests of Nominal Confidence Interval Coverage for Causal Parameters Estimated by Machine Learning”

Edward H. Kennedy, Sivaraman Balakrishnan and Larry Wasserman

REFERENCES


Edward H. Kennedy is Assistant Professor, Department of Statistics & Data Science, Carnegie Mellon University, USA (e-mail: edward@stat.cmu.edu). Sivaraman Balakrishnan is Assistant Professor, Department of Statistics & Data Science, Carnegie Mellon University, USA (e-mail: siva@stat.cmu.edu). Larry Wasserman is University Professor, Department of Statistics & Data Science, Carnegie Mellon University, USA (e-mail: larry@stat.cmu.edu).
Rejoinder: On nearly assumption-free tests of nominal confidence interval coverage for causal parameters estimated by machine learning

Lin Liu, Rajarshi Mukherjee and James M. Robins

REFERENCES


EDITORIAL POLICY

The aim of *Statistical Science* is to present the full range of contemporary statistical thought at a technical level accessible to the broad community of practitioners, teachers, researchers and students of statistics and probability.

The journal will publish discussions of methodological and theoretical topics of current interest and importance, surveys of substantive research areas with promising statistical applications, comprehensive book reviews, discussions of classic articles from the statistical literature and interviews with distinguished statisticians and probabilists.

The opinions expressed are those of the authors and do not necessarily reflect those of the editors or of the IMS.

GENERAL INFORMATION


**Submissions:** Manuscripts for *Statistical Science* should be submitted online. Authors may access the Electronic Journals Management System (EJMS) at [http://www.e-publications.org/ims/submission/](http://www.e-publications.org/ims/submission/)


**Permissions Policy.** Authorization to photocopy items for internal or personal use is granted by the Institute of Mathematical Statistics. For multiple copies or reprint permission, contact The Copyright Clearance Center, 222 Rosewood Drive, Danvers, Massachusetts 01923. Telephone (978) 750-8400. [http://www.copyright.com](http://www.copyright.com). If the permission is not found at the Copyright Clearance Center, please contact the IMS Business Office: ims@imstat.org

**Correspondence.** Mail concerning membership, subscriptions, nonreceipt claims, copyright permissions, advertising or back issues should be sent to the IMS Dues and Subscription Office, 9650 Rockville Pike, Suite L 2310, Bethesda, Maryland 20814-3998. Mail concerning submissions or editorial content should be sent to the Editor of the appropriate journal. Addresses are listed on the inside back cover. Mail concerning the production of this journal should be sent to: Patrick Kelly, IMS Production Editor, Department of Statistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6340 USA

**Individual Memberships:** [http://www.imstat.org/individual-membership/](http://www.imstat.org/individual-membership/)


INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.

IMS OFFICERS

President: Susan Murphy, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

President-Elect: Regina Y. Liu, Department of Statistics, Rutgers University, Piscataway, New Jersey 08854-8019, USA

Past President: Xiao-Li Meng, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

Executive Secretary: Edsel Peña, Department of Statistics, University of South Carolina, Columbia, South Carolina 29208-001, USA

Treasurer: Zhengjun Zhang, Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706-1510, USA

Program Secretary: Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027-5927, USA

IMS EDITORS

The Annals of Statistics. Editors: Richard J. Samworth, Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WB, UK. Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027, USA

The Annals of Applied Statistics. Editor-in-Chief: Karen Kafadar, Department of Statistics, University of Virginia, Heidelberg Institute for Theoretical Studies, Charlottesville, VA 22904-4135, USA

The Annals of Probability. Editor: Amir Dembo, Department of Statistics and Department of Mathematics, Stanford University, Stanford, California 94305, USA


Statistical Science. Editor: Sonia Petrone, Department of Decision Sciences, Università Bocconi, 20100 Milano MI, Italy

The IMS Bulletin. Editor: Vlada Limic, UMR 7501 de l’Université de Strasbourg et du CNRS, 7 rue René Descartes, 67084 Strasbourg Cedex, France
BRADLEY EFRON
TREVOR HASTIE

COMPUTER AGE
STATISTICAL
INFERENCE

ALGORITHMS, EVIDENCE, AND DATA SCIENCE