References for
Recent Developments in Cluster-Robust Inference

A. Colin Cameron

Department of Economics, U.C. Davis

Douglas L. Miller

Brooks School of Public Policy, and Economics, Cornell University

November 2022

These references accompany the November 2022 version of the slides
available at http://cameron.econ.ucdavis.edu/research/papers.html.
They emphasize more recent research.

1. Introduction

Cameron, A. Colin and Douglas L. Miller (2015), "A Practitioners Guide to Cluster-Robust Inference", *Journal of Human Resources,* 2015, 317-372.

MacKinnon, James G., Nielsen, Morten Ø. and Matthew D. Webb (2022), "Cluster-robust inference: A guide to empirical practice", *Journal of Econometrics*, in press.

2. Basics of Cluster-robust Inference

Moulton, Brent R (1986), "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics*, 32, 385-397.

Moulton, Brent R. (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics*, 72, 334-38.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (20040, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119, 249-275.

White, Halbert (1984), *Asymptotic Theory for Econometricians,* San Diego: Academic Press.

Liang, Kung-Yee, and Scott L. Zeger (1986), "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika*, 73, 13-22.

Arellano, Manuel (1987), "Computing Robust Standard Errors for Within-Group Estimators," *Oxford Bulletin of Economics and Statistics*, 49, 431-434.

Hansen, Christian (2007), "Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data when T is Large," *Journal of Econometrics*, 141, 597-620.

Carter, Andrew V., Kevin T. Schnepel, and Douglas G. Steigerwald (2017), Asymptotic Behavior of a t Test Robust to Cluster Heterogeneity," *Review of Economics and Statistics*, 99(4), 698–709.

Hansen, Bruce, and Seojeong Lee (2019), "Asymptotic Theory for Clustered Samples," *Journal of Econometrics*, 210, 268-290.

Djogbenou, Antoine, James G. MacKinnon, and Morten Ø. Nielsen (2019), "Asymptotic theory and wild bootstrap inference with clustered errors," *Journal of Econometrics*, 212, 393-412.

3. Better One-way Cluster-Robust Inference
Small-cluster bias in standard error

MacKinnon, James G., and Halbert White (1985), "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics*, 29, 305-325.

Bell, Robert M., and Daniel F. McCaffrey (2002), "Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples," *Survey Methodology*, 28, 169-179.

MacKinnon, James G., Nielsen, Morten Ø. and Matthew D. Webb (2022), "Cluster-robust inference: A guide to empirical practice", *Journal of Econometrics*, forthcoming.

MacKinnon, James G., Morten Ø. Nielsen, and Matthew D. Webb (2022), "Leverage, Influence, and the Jackknife in Clustered Regression Models: Reliable Inference Using summclust," QED WP 1483.

Hansen, Bruce E. (2022), "Jackknife Standard Errors for Clustered Regression," WP.

MacKinnon, James G., Morten Ø. Nielsen, and Matthew D. Webb (2022), "Fast and reliable jackknife and bootstrap methods for cluster-robust inference," QED Working Paper No. 1485.

Sasaki, Yuya and Yulong Wang (2022), "Non-Robustness of the Cluster-Robust Inference: with a Proposal of a New Robust Method," arXiv:2210.16991v1.

T with Different Degrees of Freedom

Imbens, Guido W., and Michal Kolesar (2016), "Robust Standard Errors in Small Samples: Some Practical Advice," *Review of Economics and Statistics*, 98, 701-712.

Pustejovsky, James E., and Elizabeth Tipton (2018),"Small sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models," *Journal of Business and Economic Statistics*, 36, 672-683.

Carter, Andrew V., Kevin T. Schnepel, and Douglas G. Steigerwald (2017), Asymptotic Behavior of a t Test Robust to Cluster Heterogeneity," *Review of Economics and Statistics*, 99(4), 698–709.

Lee, Chang Hyung and Douglas G. Steigerwald (2018), "Inference for Clustered Data," *Stata Journal*, 18(2): 447-460.

Exact Distribution other than T

Meiselman, Akiva Y. (2021), "An Exact Hypothesis Test for Samples with Few Effective Clusters," November 10, JMP UT-Austin.

Wild Cluster Bootstrap

Cameron, A. Colin, Jonah G. Gelbach, and Douglas L. Miller (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics,* 90, 414-427.

Roodman, David, James G. MacKinnon, Morten Ø. Nielsen, and Matthew D. Webb (2019), "Fast and Wild: Bootstrap Inference in Stata using boottest," *Stata Journal*, 19(1), 4-60.

MacKinnon, James G. (2022), "Fast cluster bootstrap methods for linear regression models," *Econometrics and Statistics*, forthcoming.

Webb, Matthew D. (2014), "Reworking Wild Bootstrap Based Inference for Clustered Errors," QED WP 1315.

MacKinnon, James G. and Matthew D. Webb (2017b), "Wild Bootstrap Inference for Wildly Different Cluster Sizes", *Journal of Applied Econometrics*, 32, 233–254.

Djogbenou, Antoine , James G. MacKinnon, and Morten Ø. Nielsen (2019), "Asymptotic theory and wild bootstrap inference with clustered errors," *Journal of Econometrics*, 212, 393-412.

Canay, Ivan A., Andres Santos and Azeem M. Shaikh (2021), "The Wild Bootstrap with a Small Number of Large Clusters," *Review of Economics and Statistics*, 103, 343-363.

Few Treated Clusters

MacKinnon, James G. and Matthew D. Webb (2017b), "'Pitfalls when Estimating Treatment Effects with Clustered data," *The Political Methodologist*, QED WP 1387.

MacKinnon, James G. and Matthew D. Webb (2017b), "The Wild Bootstrap for Few (Treated) Clusters," Econometrics Journal, 21, 114-135.

Ibragimov, Rustam and Ulrich K. Müller (2010), "t-Statistic based Correlation and heterogeneity Robust inference," JBES, 28, 453-468.

Ibragimov, Rustam and Ulrich K. Müller (2016), "Inference with few heterogeneous clusters", *Review of Economics and Statistics*, 98(1), 83-96.

Conley, Timothy G. and Christopher R. Taber (2011), "Inference with 'Difference in Differences' with a Small Number of Policy Changes," *Review of Economics and Statistics*, 93(1), 113-125.

Ferman, Bruno and Cristine Pinto (2019), "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity," *Review of Economics and Statistics*, 101, 452-467.

Randomization Inference

Young, Alwyn (2019), "Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results," *Quarterly Journal of Economics*, 134, 557-598.

MacKinnon, James G. and Matthew D. Webb (2020), "Randomization inference for difference-in-differences with few treated clusters", *Journal of Econometrics*, 218, 435-450.

MacKinnon, James G. and Matthew D. Webb (2019), "Wild bootstrap randomization inference for few treated clusters," in *The Econometrics of Complex Survey Data: Theory and Applications*, eds. Kim P. Huynh, David Tomás Jacho-Chávez, and Gautam Tripathi, Vol. 39 of Advances in Econometrics, Emerald Group, 2019, 61-85. QED WP 1404.

Canay, Ivan A., Joseph P. Romano, and Azeem M. Shaikh (2017), "Randomization tests under an approximate symmetry assumption," *Econometrica*, 85(3), 1013-1030.

Hagemann, Andreas (2019), "Placebo inference on treatment effects when the number of clusters is small," *Journal of Econometrics*, 213, 190-209.

Hagemann, Andreas (2020), "Inference with a single treated cluster," WP, https://hgmn.github.io/.

Hagemann, Andreas (2021), "Permutation inference with a finite number of heterogeneous clusters," WP, https://hgmn.github.io/.

Young, Alwyn (2022), "Asymptotically Robust Permutation-based Randomization Confidence Intervals for Parametric OLS Regression," WP.

Toulis, Panos (2022), "Invariant Inference via Residual Randomization," arXiv:1908.04218v2.

Design-based Inference

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge (2022), "When should you adjust standard errors for clustering?" *Quarterly Journal of Economics*, forthcoming. WP, arXiv:1710.02926v3. Revision of 2017 NBER WP.

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge, "Sampling-based versus design-based uncertainty in regression analysis," *Econometrica*, 88(1), 265–296.

Neyman, J. (1923, 1990), "On the Application of probability theory to Agricultural Experiments. Essay on Principles, Section 9," *Statistical Science*, 5, 465-472.

Startz, Richard and Dougls G. Steigerwald (2022), "The Variance of Regression Coefficients when the Population is Finite", WP, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4002003

Su, Fangzhou, and Ding Peng (2021), "Model-assisted analyses of cluster-randomized experiments," *Journal of the Royal Statistical Society* (B), 83, 994-1015.

3. Beyond Oneway Clustering
Multiway Clustering

Cameron, A. Colin, Jonah G. Gelbach, and Douglas L. Miller. (2006, 2011), "Robust Inference with Multi-Way Clustering." NBER Technical Working Paper 0327 and *Review of Economics and Statistics*. 90: 414-427.

Miglioretti, D. L., and P. J. Heagerty (2006), "Marginal Modeling of Nonnested Multilevel Data using Standard Software," *American Journal of Epidemiology* 165: 453-463.

Thompson, Samuel (2006, 2011), "Simple Formulas for Standard Errors that Cluster by Both Firm and Time." SSRN paper. http://ssrn.com/abstract=914002 and *Journal of Financial Economics* 99(1): 1-10.

Davezies, Laurent, Xavier D'Haultfoeuille and Yannick Guyonvarch (2021), "Empirical Process Results for Exchangeable Arrays," *Annals of Statistics*, 49, 845-862.

Menzel, Konrad (2021), "Bootstrap with cluster-dependence in two or more dimensions," *Econometrica*, 89, 2143–2188.

MacKinnon, James G., Morten Ø. Nielsen, and Matthew D. Webb (2021), "Wild Bootstrap and asymptotic inference with multiway clustering," *Journal of Business and Economic Statistics*, 39, 505-519.

Chiang, Harold D., Kengo Kato and Yuya Sasaki (2021), "Inference for high-dimensional exchangeable arrays," *Journal of the American Statistical Association*, forthcoming.

Chiang, Harold D., Bruce E. Hansen and Yuya Sasaki (2022), "Standard Errors for Two-Way Clustering with Serially Correlated Time Effects," WP https://www.ssc.wisc.edu/~bhansen/papers/tw

Powell, David (2020), "Panel Data Inference with Dependent Clusters," April 2020, WP.

Chiang, Harold D., Kengo Kato, Yukun Ma and Yuya Sasaki (2022), "Multiway Cluster Robust Double/debiased Machine Learning Inference for high-dimensional exchangeable arrays," *Journal of Business and Economic Statistics*, forthcoming.

Verdier, Valentin (2020), "Estimation and Inference for Linear Models with Two-way Fixed Effects and Sparsely Matched Data", *Review of Economics and Statistics*, 102(1), 1-6.

Dyadic Clustering

Fafchamps, Marcel, and Flore Gubert (2007), "The Formation of Risk Sharing Networks." *Journal of Development Economics* 83: 326-350.

Cameron, A. Colin and Douglas L. Miller (2014), "Robust Inference for Dyadic Data," WP, http://cameron.econ.ucdavis.edu/research/papers.html

Aranow, Peter M., Cyrus Samii and Valentina A. Assenova (2015), "Cluster–Robust Variance Estimation for Dyadic Data," *Political Analysis*, 23(4), 564-577.

Tabord-Meehan, Max (2019), "Inference with dyadic data: Asymptotic behavior of the dyadic-robust t-statistic," *Journal of Business and Economic Statistics*, 37(4), 671-680.

Graham, Bryan, Fengshi Niu and James L. Powell (2019), Kernel Density Estimation for undirected dyadic data," WP.

Chiang, Harold D., Yukum Ma, Joel B. Rodrigue and Yuya Sasaki (2022), "Dyadic Double/debiased Machine Learning Analyzing Determinants of Free Trade Agreements," *Journal of Business and Economic Statistics*, forthcoming.

Spatial Clustering

Driscoll, John C. and Arrt C. Kraay (1998), "Consistent Covariance Matrix Estimation with Spatially Dependent Data," *Review of Economics and Statistics*, 549-560.

Cao, Jainfei, Christian Hansen, Lucciano Villacorta and Damian Kozbur (2021), "Inference for Dependent Data with Learned Clusters," WP. arXiv preprint arXiv:2107.14677

Estimators Other than OLS
Feasible GLS

Brewer, Mike, Thomas F. Crossley and Robert Joyce (2017), "Inference with Difference-in-Differences Revisited," *Journal of Econometric Methods*, 7(1), 1-16.

Instrumental Variables

Chernozhukov, Victor, and Christian Hansen. 2008. "The Reduced Form: A Simple Approach to Inference with Weak Instruments." *Economics Letters* 100: 68-71.

Olea, Carolin and Jose Luis Montiel Olea (2013), "A Robust Test for Weak Instruments," *Journal of Business and Economic Statistics*, 31, 358-369.

Magnusson, Leandro (2010), "Inference in limited dependent variable models robust to weak identification," *Econometrics Journal*, 13, S56–S79.

Finlay, Keith and Leandro M. Magnusson (2019), "Two applications of wild bootstrap methods to improve inference in cluster-IV models," *Journal of Applied Econometrics*, 34, 911-923.

Young, Alwyn (2021), "Leverage, Heteroskedasticity and Instrumental Variables in Practical Applications," WP.

Nonlinear Estimators

Kline, Patrick, and Andres Santos (2012), "A Score Based Approach to Wild Bootstrap Inference." *Journal of Econometric Methods*:1(1): 23-41.

GMM

Hansen, Bruce, and Seojeong Lee (2019), "Asymptotic Theory for Clustered Samples," *Journal of Econometrics*, 210, 268-290.

Hansen, Bruce, and Seojeong Lee (2021), "Inference for Iterated GMM under Misspecification," *Econometrica*, 89(3), 1419-1447.

Hansen, Bruce, and Seojeong Lee (2020), "Inference for GMM under Misspecification and Clustering," WP.

Hwang, Jungbin (2019), "Simple and Trustworthy Cluster-Robust GMM Inference," *Journal of Econometrics*, 222(2), 923-1023.

Quantile Regression

Parente, Paulo M.D.C. and João M.C. Santos Silva (2016), "Quantile regression with clustered data," *Journal of Econometric Methods*, 5(1), 1-15.

Yoon, Jungmo and Galvao, Antonio F. (2020), "Cluster robust covariance matrix estimation in panel quantile regression with individual fixed effects," *Quantitative Economics*, 11(2), 579-608.

Hagemann, Andreas (2017), "Cluster-robust bootstrap inference in quantile regression models", *Journal of the American Statistical Association*, 112(517), 446-456.

Machine Learning

Cameron, A. Colin and Pravin K. Trivedi (2022), *Microeconometrics using Stata, Volume II: Nonlinear Models and Causal Inference Methods*, second edition, chapter 28.

Chernozhukov, Victor, Denis Chetverikov, M. Demirer, Ester Duflo, Christian Hansen, Whitney Newey, and James Robins, J. (2018a), "Double/debiased Machine Learning for Treatment and Structural Parameters," *Econometrics Journal*, 21, C1 – C68.

Belloni, Alex, Victor Chernozhukov, Christian Hansen, C., and Damien Kozbur (2016), "Inference in High-dimensional Panel Models With an Application to Gun Control," *Journal of Business and Economic Statistics*, 34, 590–605.

Chiang, Harold D., Kengo Kato, Yukun Ma and Yuya Sasaki (2022), "Multiway Cluster Robust Double/debiased Machine Learning Inference for high-dimensional exchangeable arrays," *Journal of Business and Economic Statistics*, forthcoming.

Chiang, Harold D., Yukum Ma, Joel B. Rodrigue and Yuya Sasaki (2022), "Dyadic Double/debiased Machine Learning Analyzing Determinants of Free Trade Agreements," *Journal of Business and Economic Statistics*, forthcoming.

Cattaneo, Matias D., Michael Jansson, and Whitney K. Newey (2018), "Inference in linear regression models with many covariates and heteroscedasticity," *Journal of the American Statistical Association*, 113, 1350–1361.

Chenchuan (Mark) Li and Ulrich K. Müller (2021a), "Linear regression with many controls of limited explanatory power," Quantitative Economics, 12, 405–442.

Chenchuan (Mark) Li and Ulrich K. Müller (2021b), "Supplement to 'Linear regression with many controls of limited explanatory power'," Quantitative Economics, 12, 405–442.

Riccardo D'Adamo (2019), "Cluster-Robust Standard Errors for Linear Regression Models with Many Controls", arXiv:1806.07314v3.