1151-62-151 **Bryan Keller***, Teachers College, Columbia University, 525 W 120th Street, New York, NY 10027. *Covariate Selection for Causal Inference based on Nonparametric Conditional Independence Testing with Random Forests.*

When randomized experimental data are available, the average causal effect of an intervention may be consistently estimated by the difference in sample means (e.g., treatment – control) because randomization guarantees balance in expectation on all observed and unobserved covariates. When randomized experimental data are not available, it may still be possible to consistently estimate an average causal effect by conditioning on observed covariates to restore balance to "as-if randomized". Rich data sets (e.g., in education or health) often yield dozens, hundreds, or more potential covariates. And while it is possible to condition on all observed covariates, it can be shown that in many cases this strategy will lead to inefficient estimation with poor finite sample properties. In this talk we review theory for covariate selection that relies on conditional independence testing and propose and implement an algorithm for nonparametric conditional independence testing based on random forests. The algorithm and a few competitors are then used for covariate selection in Monte Carlo simulations and a real-data application. Results and implications will be discussed. (Received August 15, 2019)