Abstract: There has been considerable interest across several fields in methods that reduce the problem of learning good treatment assignment policies to the problem of accurate policy evaluation. Given a class of candidate policies, these methods first effectively evaluate each policy individually, and then learn a policy by optimizing the estimated value function; such approaches are guaranteed to be risk-consistent whenever the policy value estimates are uniformly consistent. However, despite the wealth of proposed methods, the literature remains largely silent on questions of statistical efficiency: there are only limited results characterizing which policy evaluation strategies lead to better learned policies than others, or what the optimal policy evaluation strategies are. In this paper, we build on classical results in semiparametric efficiency theory to develop quasi-optimal methods for policy learning; in particular, we propose a class of policy value estimators that, when optimized, yield regret bounds for the learned policy that scale with the semiparametric efficient variance for policy evaluation. On a practical level, our result suggests new methods for policy learning motivated by semiparametric efficiency theory.

Bio: I am a postdoctoral researcher in statistics at Columbia University, and will begin an appointment as assistant professor of Operations, Information, and Technology at Stanford Graduate School of Business in the fall of 2017. I completed my PhD in statistics at Stanford University in 2016, under the supervision of Brad Efron and Guenther Walther.

My research focuses on adapting ideas from machine learning to statistical problems that arise in scientific applications. I am particularly interested in causal inference, non-parametric statistics, uses of subsampling for data analysis, and empirical Bayes methods.

I am currently serving as an associate editor for Biometrika.