

# IAB-Colloquium

## zur Arbeitsmarkt- und Berufsforschung

### A Comparison of Machine Learning Techniques for Modeling Survey Nonresponse

**Prof. Curtis S. Signorino**

*University of Rochester, New York*

Increasing reluctance to participate in surveys leads to low response rates – sometimes as low as 20-30%. It is well known that when (non)response is not completely at random, estimates calculated based on the resulting sample may be severely biased. When auxiliary information is available on both respondents and nonrespondents, one common method for correcting nonresponse bias is to model response propensity and then weight responses by the inverse of their response probabilities. Typically, a logit or probit with a linear link is used to model response propensity. Recent developments in machine learning allow for flexible functional form estimation and variable selection. In this project, we use Adaptive LASSO applied to basis regression – here a polynomial expansion of the regressors. Adaptive LASSO is one of a few recent penalized (or shrinkage) estimators. These techniques are especially useful when there are many regressors, only a few of which may be relevant, because they shrink coefficients of irrelevant variables to zero and have impressive consistency properties. Kenkel & Signorino (2012, 2013) have shown this technique (implemented in the R package polywog) to be effective at modeling flexible functional forms that result from nonrandom selection or from strategic choice, even when many irrelevant (and possibly correlated) regressors are included in the regression model. We apply this technique to modeling response propensities for the panel study ‘Labour Market and Social Security (PASS)’, using administrative data for both respondents as well as nonrespondents. We compare the nonresponse bias correction when using Adaptive LASSO with a polynomial expansion of regressors to existing techniques: logistic regression, neural nets, classification trees, and random forests.

**Dienstag, 12. August 2014**

13:00 Uhr

Sitzungssaal 126a