Letter to the Editor: Treatment prediction, balance and propensity score adjustment: Supplementary figures

Erica E. M. Moodie and David A. Stephens

Additional figures indicating the overlap in propensity score distribution by estimation method (eFigure 1), balance in the distribution of age (eFigure 2), and a summary of standardized mean differences for interactions by estimation method (eFigure 3) are provided below.

Figure 1. Boxplots of the propensity score distribution in non-smokers (0) and smokers (1), with red horizontal lines marking quintiles of the distribution. Left to right: propensity score estimated by logistic regression, generalized boosted models, and Super Learning.
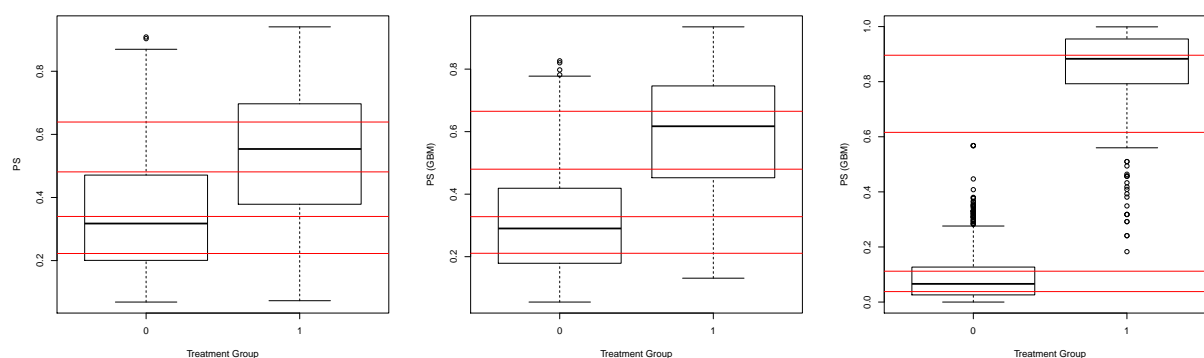
Figure 2. A visual check of balance in the continuous covariate, age: empirical cumulative distribution plots for non-smokers (black) and smokers (red). Top to bottom: original sample, and propensity score estimated by logistic regression, generalized boosted models, and Super Learning. Left to right, rows 2 to 4: propensity score matched and inverse probability weighted.
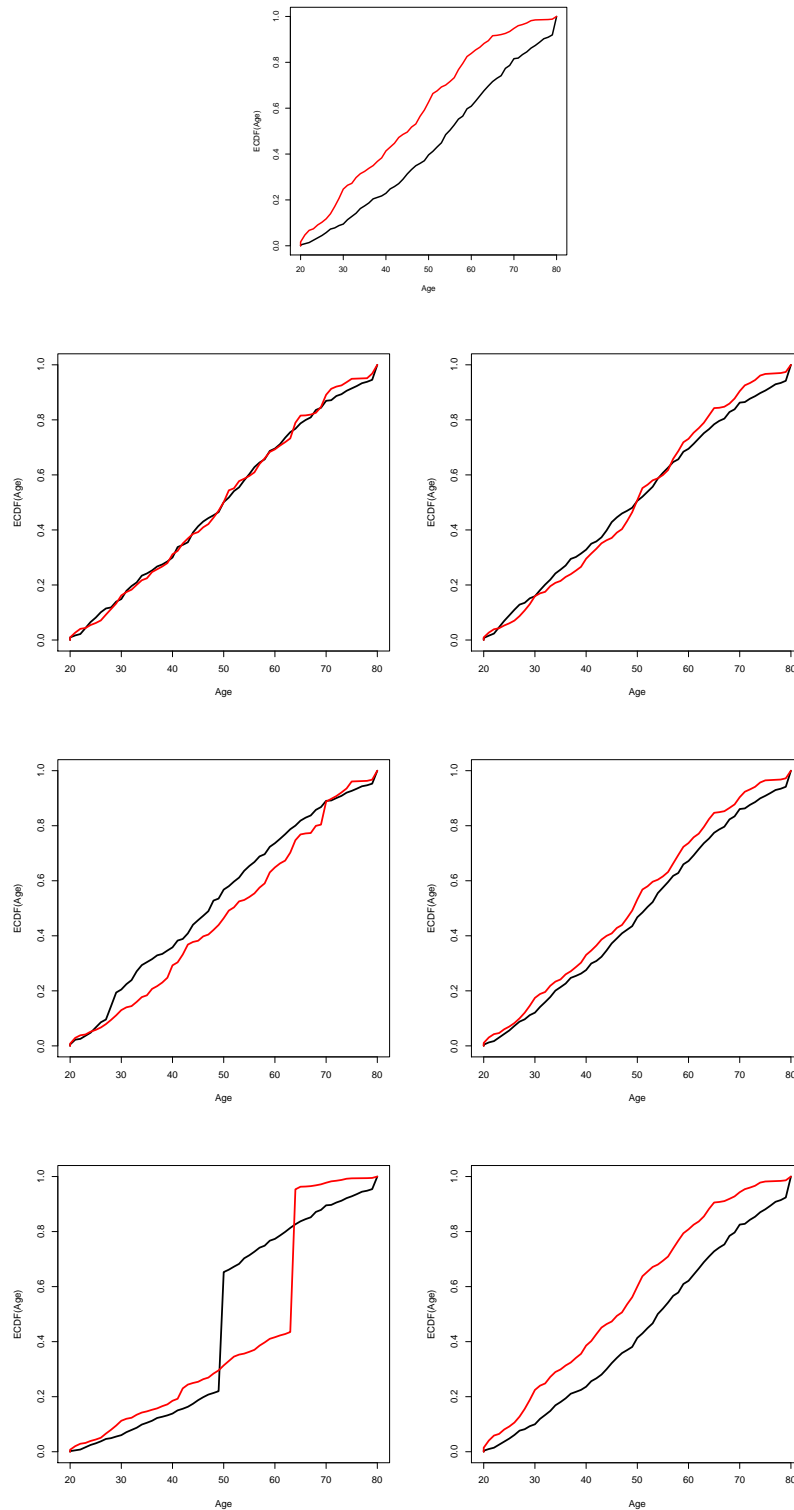
To better assess whether balance might hold jointly across covariates for each of the methods of estimation, we began by considering all possible first-order interactions. We removed the interaction for any covariate pair for which data were sparse, which we defined as having five or fewer individuals in any level of the covariate-pair combination. This left 14 pair-wise interactions, which, if coded as covariates in their own right, yielded 107 terms on which balance could be assessed. As inverse probability weighting showed the best balance across all three methods, we consider the balance of these 107 variables in the inverse probability of treatment weighted samples; the distribution of the standardized mean difference in the original sample and the three weighted samples (one per method of estimating the treatment model) are plotted in eFigure 3 below.

Figure 3. Boxplots of the standardized mean differences for interaction terms in the original (unadjusted) sample and in inverse probability weighted (IPW) samples by estimation method. Left to right: original sample, and propensity score estimated by logistic regression (LR), generalized boosted models (GBM), and Super Learning (SL).