American Political Science Review (2023) 117, 2, 785–787
doi:10.1017/S0003055422001435

Corrigendum

BARP: Improving Mister P Using Bayesian Additive Regression Trees — CORRIGENDUM

MAX GOPLERUD *University of Pittsburgh, United States*
JAMES BISBEE *Vanderbilt University, United States*

DOI: https://doi.org/10.1017/S0003055419000480, Published by Cambridge University Press, 06 August 2019.

Bisbee regrets the coding error in the above article. As identified by Goplerud (Forthcoming) [Appendix E.1], Bisbee (2019a)'s replication code failed to sort the data prior to calculating predictions from the MRP model, leading to the injection of noise into MRP estimates while not affecting BARP estimates. This lead to exaggerated performance improvements when comparing traditional MRP with BARP. The error affects Figures 1, 2, and 3 in the original publication as well as figures in the supplementary materials. Bisbee has corrected figures for the main text which are reproduced below (similar figures appear in Goplerud Forthcoming), and he has updated the associated supporting materials for Bisbee (2019a) to reflect this correction.

The corrected version of Figure 1 and its caption are provided below. The corrected results demonstrate that the difference in performance between the two methods is much more of a toss-up, whether evaluated using mean absolute error (MAE, left panel) or interstate correlation (right panel). Goplerud (Forthcoming) summarizes the difference quantitatively by averaging across the surveys and reports a small improvement of BARP over (traditional) MRP of around 4.5% with a sample size of 1,500. When considering mean absolute error, it notes that this decreases to around 1% for larger sample sizes.

The corrected version of Figure 2 and its caption are provided below. It also shows a much more similar performance between the two methods across the range of surveys considered.

The corrected version of Figure 3 and its caption are provided below. It shows results that are consistent with the claim in Bisbee (2019a), i.e. that BARP is less sensitive to smaller sample sizes. Appendix E.2 of Goplerud (Forthcoming) provides a different test of this claim and finds limited differences between the methods.

CONCLUSION

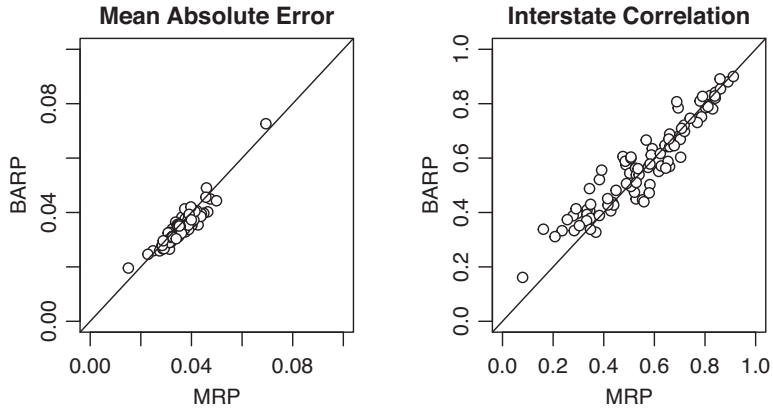
After the discovery the original error, the two authors of this corrigendum spoke, agreed on the source and nature of the error, and then jointly wrote this correction. Bisbee has re-examined the associated software for implementing BART for MRP and confirmed that the error did not affect recent research that has relied on this software.[1] Updated replication materials are available at Bisbee (2019b).

REFERENCES

Bisbee, James. 2019a. "BARP: Improving Mister P Using Bayesian Additive Regression Trees." *American Political Science Review* 113(4): 1060–5.
Bisbee, James. 2019b. "Replication Data for: BARP: Improving Mister P Using Bayesian Additive Regression Trees, V2." Harvard Dataverse. Dataset. https://doi.org/10.7910/DVN/LMW871.
Goplerud, Max. Forthcoming. "Re-Evaluating Machine Learning for MRP Given the Comparable Performance of (Deep) Hierarchical Models." *American Political Science Review*.

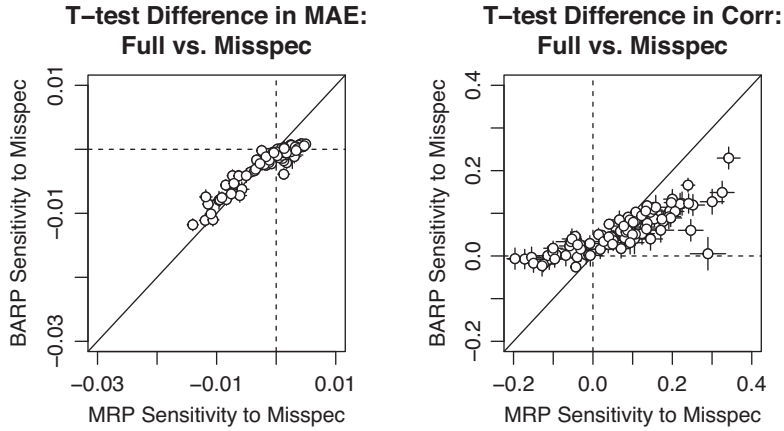
[1] https://github.com/jbisbee1/BARP

FIGURE 1. Predictive Accuracy



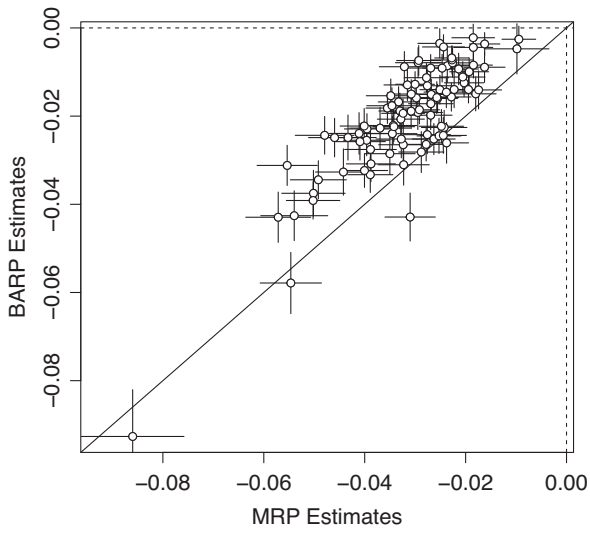
Note: Predictive accuracy of BARP (y-axes) versus MRP (x-axes) across 89 surveys as measured by mean absolute error (left panel) and interstate correlation (right panel).

FIGURE 2. Sensitivity to Misspecification



Note: Difference-in-means estimates (points) and confidence intervals (lines) indicating how much better MRP (x-axes) and BARP (y-axes) perform when the two state-level covariates are included. Negative values on the left-hand plot reflect smaller absolute errors in the full specification, whereas positive values on the right-hand plot reflect larger interstate correlations in the full specification.

FIGURE 3. Method Sensitivity to State Sample Size



Note: Coefficients (points) for each survey measuring the relationship between mean absolute error and the number of observations in the state for BARP (y-axis) and MRP (x-axis). Negative values indicate that more observations in a state improve improve mean absolute error by the units on the x and y-axes. Two standard errors indicated by horizontal and vertical lines. Values closer to zero (dashed lines) reflect greater insulation from data sparsity.