Statistical Science
2020, Vol. 35, No. 3, 371–374
https://doi.org/10.1214/20-STS790
Main article: https://doi.org/10.1214/19-STS699

Rejoinder: Matching Methods for Observational Studies Derived from Large Administrative Databases

Ruoqi Yu, Jeffrey H. Silber and Paul R. Rosenbaum

1. OUTLINE

We thank the discussants for their insightful and generous comments. We organized our reply around a few themes, rather than responding to issues one by one. In Section 2, we recap the major elements of the paper in light of the discussion. Then Section 3 reviews the several goals of matching. Finally, Section 4 discusses open questions.

2. RECAP

First, let us restate the main themes of the paper.

- *Network optimization*. In our paper, each matched sample is obtained by optimizing a criterion subject to constraints. Specifically, each match is obtained as a minimum cost flow in a network, a rich but special family of integer programs that can be solved in polynomial time; see Bertsekas (1998) and Korte and Vygen (2012). There are other approaches to matching that leave the world of polynomial-time optimization algorithms, and these have both advantages and disadvantages (Zubizarreta, 2012; Karmakar, Small and Rosenbaum, 2019), but they are not discussed in our paper.
- *The constraints do most of the covariate balancing*. It is not possible to closely pair individuals for many covariates. It is possible to form treated and control groups with similar distributions of many covariates; that is, it is possible to balance many low-dimensional summaries of high-dimensional covariates. Balancing of covariates is largely achieved by the constraints, not by minimizing the within pair covariate distance. The balancing constraints include: (i) calipers on the rank of the scalar propensity score, (ii) near-fine balance constraints for a nominal covariate, perhaps with thousands of levels, (iii) possibly other balance constraints (Zubizarreta, 2012; Yu and Rosenbaum, 2019). If the constraints do most of the work, then finding the constraints that achieve your objectives is a central aspect of matching. In contrast, covariate distances used for pairing should focus on a few key covariates highly predictive of the outcome (Rosenbaum, 2005; Zubizarreta, Paredes and Rosenbaum, 2014).
- *Optimization is not recommendation*. As our example illustrates, the standard practice is to build several optimal matched samples, then pick the best one. There is no contradiction here: an optimal match is the solution to an optimization problem, not a recommended match. The practical goal is a match that is good in several senses, not best in one overriding sense, so each optimal solution is merely an approximation to the practical goal. Optimization is an aid to judgement, not a substitute for judgement. It is possible to produce the set of Pareto optima for a multi-objective optimization problem as a potentially useful guide (Pimentel and Kelz, 2020; Rosenbaum, 2012), but ultimately the investigators must pick one match, so the basic structure is unchanged: practical judgement is used to pick the most satisfactory of several optimally matched samples. Several optimal solutions provide points on a map by which judgement can steer among multiple objectives. Matching is part of the design of the study, completed prior to the examination of outcomes (Rubin, 2007).
- *Guarantee feasibility; guarantee speed*. There is no point in trying to solve an optimization problem subject to constraints if no solution satisfies the constraints. A fast implementation of Glover's (1967) algorithm permits certain types of constraints to be checked for feasibility at negligible cost. These include combinations of: (i) exact match constraints for a nominal covariate, perhaps with many levels, (ii) a caliper on the rank of the propensity score, (iii) a near-neighbor count on the propensity score. A threshold algorithm—a binary search—rapidly finds the tightest feasible constraint with negligible error, thereby guiding optimal

Ruoqi Yu is PhD student, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6340, USA (e-mail: ruoqiyu@wharton.upenn.edu). Jeffrey H. Silber is Professor, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA (e-mail: silber@email.chop.edu; URL: https://cor.research.chop.edu/). Paul R. Rosenbaum is Professor, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6340, USA (e-mail: rosenbaum@wharton.upenn.edu).