Large, Sparse Optimal Matching in an Observational Study of Surgical Outcomes

Abstract: Every newly trained surgeon performs a first unsupervised operation. How do her patients' health outcomes compare with the patients of experienced surgeons? A credible comparison must (1) occur within hospitals, since health outcomes vary widely by hospital; (2) compare outcomes of patients undergoing the same operative procedures, since the risks differ in a knee replacement and an appendectomy; (3) control for potentially higher risks among patients from the emergency room or from a less capable hospital, since new surgeons treat them at a disproportionate rate; and (4) compare patient samples with similar distributions of health problems such as diabetes. We introduce a new form of matching that pairs patients of 1252 new surgeons to patients of experienced surgeons, exactly balancing 176 surgical procedures and closely balancing 2.9 million finer patient categories. The new matching algorithm (which uses penalized network flows) exploits a sparse network to quickly optimize a match two orders of magnitude larger than usual in statistical matching. This allows extensive use of a new form of marginal balance constraint called refined covariate balance. The match was constructed in minutes using software in R.