Abstract

Student evaluations of teaching are widely used to measure teaching quality and compare it across different courses, teachers, departments and institutions: as such, they are of increasing importance for teacher promotion decisions as well as student course selection. However, the response on course evaluations is rarely perfect, rendering such uses unwarranted if students who participate in the evaluation are not randomly selected: this paper is the first to investigate this issue. We quantify the direction and size of selection on both observable and unobservable characteristics for a large European university where course evaluation response rates differ across courses. Our results suggest course evaluations are upward biased, and that this bias mostly derives from selection on characteristics unlikely to be observed by the typical university administrator. Correcting for selection bias has sizable effects on both scores in any given course and the evaluation-based ranking of different courses.