Student Evaluations, Quantifauxcation, and Gender Bias

Abstract:
Student evaluations of teaching (SET) are widely used in academic personnel decisions as a measure of teaching effectiveness. The way SET are used is statistically unsound—but worse, SET are biased and unreliable. Observational evidence shows that student ratings vary with instructor gender, ethnicity, and attractiveness; with course rigor, mathematical content, and format; and with students’ grade expectations. Experiments show that the majority of student responses to some objective questions can be demonstrably false. A recent randomized experiment shows that giving students cookies increases SET scores. Randomized experiments show that SET are negatively associated with objective measures of teaching effectiveness and biased against female instructors by an amount that can cause more effective female instructors to get lower SET than less effective male instructors. Gender bias also affects how students rate objective aspects of teaching. It is not possible to adjust for the bias, because it depends on many factors, including course topic and student gender. Students are uniquely situated to observe some aspects of teaching and students’ opinions matter. But for the purposes of evaluating and improving teaching quality, SET are biased, unreliable, and subject to strategic manipulation. Reliance on SET for employment decisions disadvantages protected groups and may violate federal law. For some administrators, risk mitigation might be a more persuasive argument than equity for ending the use of SET in employment decisions: union arbitration and civil litigation over institutional reliance on SET are on the rise. Several major universities in the U.S. and Canada have already de-emphasized, substantially re-worked, or abandoned SET for personnel decisions.

Special Note:
Refreshments will be provided. Additional post-seminar discussion time from 2-3pm.

Organizer: CAMSEE

Monday, February 4, 2019
1:00 PM
Tata Hall - Kavli Auditorium 3201