

Stanford University
Department of Statistics

DEPARTMENTAL SEMINAR

The Stanford/Berkeley Joint Statistics Colloquium

4:30pm, Tuesday, October 30, 2018

*** Venue Change ***

Jordan Hall Building 420 Room 041

Refreshments served at 4pm in Sequoia Lounge.

Speaker: Philip Stark, *UC Berkeley*

Title: **Student Evaluations of Teaching Do Not Measure
Teaching Effectiveness: What Do They Measure?**

Abstract:

Student evaluations of teaching (SET) are widely used in academic personnel decisions as a measure of teaching effectiveness. 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. Moreover, student responses to objective questions, such as the timeliness with which assignments are returned, are often demonstrably false. 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 affects how students rate even "objective" aspects of teaching, such as how promptly assignments are graded. 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 they should be consulted, 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. Union arbitration and civil litigation over institutional use of SET are on the rise, and a number of universities are de-emphasizing or abandoning SET for employment decisions.