What They Say and What They Mean:
Modeling Misreported Counts
Daniel Heitjan
Professor, Department of Statistical Science
Southern Methodist University
Professor of Clinical Sciences
UT Southwestern Medical Center
January 31, 2019
3:30 p.m.
Marrs McLean Science Building 101

Abstract: Self-reported count data are subject to a variety of types of measurement error. One distinctive form is what we call “heaping,” or the tendency of subjects to report round numbers in place of true values. Heaping appears in many types of data—including ages, numbers of cigarettes smoked, numbers of sexual partners, and church attendance.

Heaping differs fundamentally from other types of measurement error, and therefore is not amenable to treatment with standard statistical methods for measurement error. Statisticians have made progress on analyzing heaped data by considering it a form of missing or incomplete data. But such methods rely on specification of a “coarsening mechanism,” meaning a stochastic model describing how one gets from the true underlying data value (“what they mean”) to the observed value (“what they say”). Model parameter estimates can be sensitive to this specification, however; thus when the true mechanism is not known, answers from putative models are unreliable.

In this talk, I will describe a unique data set that includes both true and reported values of daily cigarette counts, enabling precise estimation of the coarsening mechanism. I will present the results of several model-based statistical analyses of the data. Findings suggest that cigarette count data are perhaps less accurate than we had believed, and that models that invoke “rounding” as an explanation for the heaps are also inaccurate.

Daniel Heitjan earned his PhD in Statistics with Donald Rubin at the University of Chicago in 1985. He has since served on the faculties of UCLA, Penn State, Columbia, and the University of Pennsylvania. He is currently Professor of Statistical Science at SMU and Professor of Clinical Sciences at the UT Southwestern Medical Center. In addition to teaching and research, he directs the biostatistics PhD program that his two institutions collaborate in operating. Dr. Heitjan is an elected fellow of the American Statistical Association, the Institute of Mathematical Statistics, and the Society of Clinical Trials, and recipient of the 2017 Don Owen Award from the ASA San Antonio Chapter.

Please join us for refreshments in MMSCI 179 at 3:00.