Why 'Multiply' Impute When You Can 'Sufficiently' Perturb?
A Comparison of Techniques For Protecting Confidential Numerical Data

Krish Muralidhar

University Of Kentucky, Lexington, KY 40506-0034
(859)-257-9149
krishm@uky.edu

Rathindra Sarathy

Oklahoma State University, Stillwater, OK 74078.
(405)-744-8646
sarathy@okstate.edu

ABSTRACT

In a 1993 special issue of the Journal of Official Statistics (Vol. 9, No. 2) focusing on disclosure limitation, Rubin suggested that multiple imputation might be a better method for masking microdata than perturbation methods that existed at that time. Rubin's article formed the basis for future developments of multiple imputation. In that same issue, Fuller (1993) also provided an excellent discussion of perturbation, especially as it related to numerical microdata. Fuller's paper provided the impetus for further developments in perturbation techniques. In this study, we argue that the recent advances in perturbation techniques, especially for numerical microdata, makes perturbation a superior alternative to the multiple imputation. We compare the performance of recent perturbation methods with multiple imputation using simulated data. We show that, for masking continuous numerical data, the performance of perturbation dominates that of multiple imputation. Based on this, we suggest further research to extend the recent developments in perturbation for masking numerical microdata to discrete and categorical data.

Due to space restrictions, we are unable to include a copy of the entire paper. For a copy of the complete paper, please contact the first author.