

ACMS Statistics Seminar

Claire Bowen

Urban Institute

Tuesday, September 19, 2023

101A Crowley Hall

3:30 PM – 4:30 PM



Safe Data Technologies Project: Safely Expanding Access to Administrative Tax Data

The Statistics of Income (SOI) Division within the Internal Revenue Service curates and maintains an extensive repository of tax-related data, offering researchers a valuable resource for evaluating the impacts of tax policies and exploring diverse research inquiries, including the analysis of income inequality. While the confidential data remains accessible only to a limited number of government analysts and researchers, the SOI provides an accessible public use file for external researchers and data practitioners. However, safeguarding this public use file has grown increasingly difficult to protect through traditional statistical data privacy methods, as the vast amount of personal information available in public and private databases combined with enormous computational power create unprecedented privacy risks.

This presentation delves into the collaborative efforts of the SOI Division and researchers at the Urban Institute, who are actively developing a solution: the creation of synthetic data that represent the statistical properties of the administrative data without revealing any individual taxpayer information. In addition to this, Urban is building a prototype validation server that empowers researchers to indirectly conduct statistical analyses on administrative tax data. Researchers can accomplish this by evaluating their analyses using synthetic data and subsequently submitting them to the validation server. The server then produces a modified output with added noise, all the while maintaining the confidentiality of taxpayer information.

In my talk, I will delve into the lessons learned, best practices, and challenges during the process of safely expanding access to administrative tax data.

The Department of Applied and Computational
Mathematics and Statistics

Please visit acms.nd.edu to view the full list of speakers.