

Title: Recent Confidentiality Research Related to Access to Enterprise Microdata

Author: Arnold Reznick, U.S. Census Bureau

This paper gives an overview of several recent developments in Statistical DisclosureControl (SDC) research and how they may increase researchers' ability to analyze confidential data. Recent basic results are starting to clarify situations in which disclosure risks may exist in regression-type models. Agencies are developing "model servers," which allow researchers to estimate statistical models from outside the agency without having direct access to the underlying data. Researchers are showing how agencies might combine confidential data sets for analysis while maintaining confidentiality of all the data sets. Perhaps most significantly, researchers are now constructing synthetic microdata sets, which mimic the properties of the underlying data but which can be accessed directly outside the agency.