NSF grant and new technology help CU-led project mine census data

By Linda Myers

A gold mine of information collected by the U.S. Bureau of the Census but previously inaccessible to researchers could be used to tackle a range of social issues, according to John M. Abowd, professor of labor economics in Cornell's School of Industrial and Labor Relations. "The new knowledge that can be generated from these data is potentially far reaching," he said, "covering everything from where the best jobs are to how to make public policy more responsive to people's needs."


Abowd

Now a $4.1 million grant from the National Science Foundation is helping Abowd and his colleagues, Julia Lane at the Urban Institute and John Haltiwanger at the University of Maryland, harness new technologies to link disparate data bases and protect the confidentiality of the underlying data, while enabling economic and other researchers to use them.

Abowd, director of the Cornell Institute for Social and Economic Research (CISER), is the principal investigator of a five-year project, "Dynamic Employer-Household Data and the Social Data Infrastructure," that was among six proposals funded by the NSF under its "Enhancing Infrastructure for the Social and Behavioral Sciences" initiative. The project is co-sponsored by the U.S. Bureau of the Census.

"These data, which could describe the dynamic interactions of workers, business, government and society, are not fully used in the United States today," said Abowd. "The grant represents a major research opportunity for the Census Bureau and social scientists across the country, who have not had access to this national statistical resource." Society as a whole is likely to reap the benefits of their research through improved quality of the Census Bureau's household and economic surveys and through better, more timely research on social issues like welfare-to-work and aging.

The development of sophisticated networking and teleconferencing technologies and high-speed computer applications make such a project possible for the first time, noted Abowd. "We now have the tools to integrate individual databases that include longitudinal data" -- information gathered over time -- "on American social and economic life, while protecting people's essential privacy."

Because the Census Bureau guarantees citizens that the information it collects from them will remain confidential, the project will create "a firewall within a firewall within a firewall," ensuring "rigorous confidentiality safeguards," said Abowd. A firewall is a computer system that protects networks against unauthorized intrusion. "The National Science Foundation has funded this research so that the Census Bureau and other national statistical agencies can be convinced that they can expand access for researchers without compromising the confidentiality of their data," Abowd added.

Before allowing researchers to use carefully constructed extracts of the data, Abowd and his colleagues will remove all personally identifying information. In addition, they will repeatedly test the effectiveness of their multiple firewalls using simulated information as well as restricted-access household and economic information lent from European statistical agencies. The project also may serve as a model for other kinds of interdisciplinary research.

March 9, 2000

| Cornell Chronicle Front Page | | Table of Contents | | Cornell News Service Home Page |