

Benefits and challenges of using administrative data for statistical purposes

Robert M. Goerge

May 10, 2019

National Statistics for Public Policy: Linkages among Federal, State and Local Data

Committee on National Statistics

Outline of presentation

- Origins of this work
- Benefits
- States Needs Assessments
- Barriers
- Addressing Barriers

Chapin Hall at the University of Chicago

- Organization that was brought to the University in order to have an impact on the well-being of children through rigorous research and analysis
- Over the past 30 years, we have built relationships with policymakers in order to have that impact through conducting research that meets their needs
- Known for our work with social program administrative data, we employ a full range of methods and address many issues of children, families and their neighborhoods

Problem Statement

- State government leaders do not have adequate information to make some of the highest stakes decisions in Illinois
 - Yet, the State has the raw data
 - But, the State lacks the processes for combining and the tools for analyzing the data across programs
 - And, the State lacks the expertise to conduct the analysis.
- The IDB Project's goal is to improve the information that is available to Illinois leadership, by improving both the data and the resources available to state and other analysts

Benefits

- Relative to primary data collection
 - Data belongs to the organization that collects it
 - They don't have to share it, except with the federal gov't sometimes
- Universe of a specific population of interest
- Longitudinal data within constraints of level of government
- Identifiers are available, thereby able to link to other data with identifiers
- Speaks directly to the daily operations of public agencies
 - For example, it's hard to collect primary data collection around abuse and neglect investigations, while administrative data captures them well
 - National surveys are challenged to capture the diversity of policy and program within states

Needs Assessment I

- Needs assessment found that states had:
 - limited analytic capacity
 - challenges around linking data, and
 - challenges around sharing data
- Allard, S. W., Wiegand, E. R., Schlecht, C., Datta, A. R., Goerge, R. M., & Weigensberg, E. (2018). State agencies' use of administrative data for improved practices: Needs, challenges, and opportunities. *Public Administration Review*, 78(2), 240-250.

Practitioner Points

- Improving data quality tends to involve intentional, but incremental, efforts to enhance the accuracy of data entry, reduce the prevalence of missing data, and create detailed data documentation.
- High-capacity administrative data users allocate staff time to reflect upon questions or metrics that extend beyond mandatory reports and will have immediate benefits to the agency and its clients.
- Agencies making consistent commitments to analytic activity outside of mandatory reporting do so by prioritizing and protecting staff time for such activities and developing data infrastructures that make data flexibly available for analytic work.
- Integrating analytics into program planning and design often requires intentional partnerships with external researchers, agencies in other states, charitable philanthropy, and federal agencies.
- States seeking to become field-leading learners build and share comprehensive data resources, participate in creative discussions about data analytics and program design, and support the learning of data users.

Needs Assessment II

- An activity the TANF Data Innovation project, funded by HHS
- Structured survey TANF program of all states and territories (N=54)
- Up to 8 respondents in each state
- Major early finding from 48 states is that they lack staff and staff with data analytic skills
- Primary interest in employment outcomes and evaluating effectiveness of their programs

Data used for compliance not an asset for states

- For many programs, the federal government requires states to submit microdata on participants, services, transactions and benefits.
- For example ... TANF, Medicaid, SNAP, foster care and adoption, abuse and neglect investigations, child care subsidies, employment and New Hires (National Directory of New Hires), vital statistics
- These data, which are provided in a common format used by all states, are seldom used for research and analysis by states
- Because
 - they do not do much analysis, in general
 - the format does not facilitate answering their questions;
 - compliance activities are separated.

Family Self-Sufficiency Data Center

- HHS-funded project increase the use and quality of administrative data in programs intended to promote family self-sufficiency
- TANF, SNAP, CCDF (Child care subsidy)
- Worked closely with ten states and engage more in workshops
- Multi-state project on employment outcomes of TANF cases
- Created Data Model and shared code to create spells of TANF receipt
- <https://chapinhall.github.io/FSSDC/>

Administrative Data Research Facility (ADRF)

- Coleridge Initiative out of NYU led by Julia Lane. Chapin Hall is a partner
- <https://cusp.nyu.edu/coleridge-initiative/>
- Provides a secure platform to host confidential micro-data.
- The ADRF system has been granted FedRAMP moderate pre-authorization and is designed to scale to different amounts and types of use.
- It provides a set of data analysis components that can be combined in different ways, within the secure ADRF boundary, to meet a wide range of analytical needs.
- *It provides a place where data can be shared across states, which for many outcomes (e.g. employment), is critical. This has been the benefit of the bringing data to the Census Bureau in the past, which seems to be getting harder.*

Overcoming the barriers

- Engage the leadership of the state agencies
- Make the case as to how the data use is beneficial for the *agency providing the data* (requires understanding their goals)
- Build a partnership based on shared objectives around the production of evidence and statistics
- **Share code to build, link, and analyze administrative data**
- **Protect data and avoid disclosure**