

CALL FOR ARTICLES

RSF: THE RUSSELL SAGE FOUNDATION JOURNAL OF THE SOCIAL SCIENCES

Issue and Conference on Using Administrative Data for Science and Policy

Andrew M. Penner

University of California, Irvine

Kenneth A. Dodge

Duke University

Administrative data sources play an increasingly central role in understanding inequality, and recent initiatives like the Murray-Ryan Evidence-Based Policymaking Commission Act of 2016 suggest that administrative data infrastructure will only become more central in the future of social science and policy. Efforts to leverage administrative data in the social sciences to understand inequality and poverty are, however, uneven: In some domains, administrative data are used routinely, while they are virtually never used in others. The quality of these data has increased greatly, particularly in education and healthcare, due to accountability requirements. The potential for linking administrative data files across domains (e.g., education data and social services data) has improved with the advent of common identifiers. This issue of *RSF: The Russell Sage Foundation Journal of the Social Sciences* seeks to highlight the promise of analyzing administrative data for understanding issues around social, political, and economic inequalities, showcasing the unique insights that such data can provide in understanding the causes and consequences of these inequalities, and the effectiveness of programs and policies aimed at redressing these.

We welcome contributions from a wide-range of disciplines and perspectives using administrative data, including (but not limited to) public health, education, economics, psychology, political science, and sociology. Whereas administrative data have been used extensively in program and policy evaluation, we are also open to rigorous descriptive research using administrative data to help us understand better the contours of inequality, to integrate qualitative and quantitative data, and to advance theory. We welcome studies using administrative data from single geographical districts or organizations as well as the entire United States. Recognizing that the potential for insight grows exponentially as data are integrated, we are particularly interested in papers that link data sources that are often siloed.

Note that while much important work on administrative data has a non-US focus, per the Russell Sage Foundation's charter, we consider work that focuses on "the improvement of social and living conditions in the United States." Fortunately, the field is rich with domestic data.

Below we offer (non-exhaustive) examples illustrative of the kinds of topics abstracts might consider.

Prospective studies of long-term outcomes. Recent research has highlighted the utility of administrative data for understanding the long-term outcomes associated with a variety of interventions. While much policy is necessarily driven by research and evaluations examining relatively short-term outcomes, administrative data can provide a longer perspective on the effects of policies, and in so doing provide a fuller account of their costs and benefits. Beyond following individuals over time, administrative data can provide an important tool for understanding multigenerational cycles of advantage and disadvantage, allowing researchers to trace the descendants of individuals from different backgrounds, as well as the multigenerational effects of anti-poverty policies and interventions. Administrative data also afford the opportunity to relate directly observed variables early in life (e.g., interview responses, school classroom behaviors) to future administrative records, so that validity can be assessed and the fuller implications of observed variables can be realized (e.g., shadow prices of school behavior problems).

Combining administrative datasets across different domains. By virtue of how they come into existence, administrative data are typically focused on one facet of an individual's life, and data and insights are often siloed. By combining administrative data from domains like schools, criminal justice institutions, health organizations, and employers, researchers can address important questions about how inequalities compound across these domains. For example, how do disciplinary actions at schools lead to later criminal justice system involvement? How do inequalities at work affect health? Likewise, combining administrative data from private companies with large national administrative datasets could help us to better understand the implications of the gig economy for workers.

Combining survey data with existing administrative records. While administrative records hold great promise for addressing vital questions, the data they include often do not measure the constructs most germane to testing theory. This limits the range of questions that analyses drawing solely on administrative data can examine. We are thus interested in studies where researchers have matched data containing researcher-fielded surveys with existing administrative records to expand the range of questions that can be addressed. Examples might include linking questions about motivational psychology to educational administrative data, linking information about vaccination attitudes with administrative records on public health, or linking information about implicit biases with administrative records containing behavioral measures of disparities (e.g., police officers' arrest decisions, managers' hiring and promotion decisions, doctors' time spent with patients).

Understanding individuals in their social contexts. Research using administrative data can provide a more complete picture of certain aspects of individuals' lives. For example, it is

difficult to imagine survey data tracking all of the classmates that a student had, or all of the co-workers over the career of an employee, but educational administrative data files and linked employer-employee datasets might include this information. Comprehensive administrative records provide information not only about a research subject but also about the environment surrounding that participant (e.g., school records provide information about a student's academic performance as well as the performance of classroom peers). Further, administrative data allow for individuals to be placed in a multi-generational familial context. In providing dense coverage of populations, these data allow researchers to examine whether policies had spillover effects on those around the targeted populations (either positive or negative), and to examine questions around treatment heterogeneity in the effects of community-level interventions (as well as other issues relevant for scaling up interventions to operate on a broader societal level).

Small but theoretically important populations. Administrative data can allow us to understand small, often difficult to access populations that are theoretically important. For example, large administrative datasets can help us understand the unique dynamics of small racial and ethnic groups, compare individuals to others who work in the same firm or establishment, and examine questions around elites like the 1% (or 0.1%). Insofar as many large administrative data sets contain information on the whole population, these data allow researchers to examine small and theoretically important groups without compromising representativeness.

Key sites in the production of inequality. Inequality is often produced in spaces that are difficult to examine using surveys or experiments. In the hiring process, for example, correspondence studies can examine who receives call backs, but cannot help us understand how applicant pools are created, which interviewees receive offers, or pay differences among those who receive offers. Research using administrative data on hiring pipelines, university admissions committees, and other key sites where gatekeepers make decisions with important consequences for inequality can thus play an important role in helping understand how inequality is produced.

Qualitative research making use of rich administrative data resources. Although much of the research using administrative data examines takes a quantitative approach, we welcome data using qualitative analyses of administrative records. For example, applicants to universities and firms often provide detailed dossiers, and archives of job postings can be used to provide insights into what employers view as important.

Highlighting the role of administrative data in the iterative policy design, implementation, and evaluation cycle. The research-policy link is often conceptualized as one in which research informs or evaluates policies, but in research-practitioner partnerships, policy implementation and research can have a bi-directional synergistic relationship. Longer term research projects on social, political, and economic inequalities highlighting the role of administrative data in an iterative process of policy design, implementation, and evaluation over time can thus provide a model for how to accumulate and incorporate knowledge rather than providing a single policy evaluation.

Anticipated Timeline

Prospective contributors should submit an abstract (up to two pages in length, single or double spaced) of their study along with up to two pages of supporting material (e.g. tables, figures, pictures, etc.) no later than **5 PM EST on June 15, 2017** to:

<https://rsfjournal.onlineapplicationportal.com>.

All submissions must be original work that has not been previously published in part or in full. Only abstracts submitted to <https://rsfjournal.onlineapplicationportal.com> will be considered. Each paper will receive a \$1,000 honorarium when the issue is published. The journal issue is being edited by Andrew M. Penner, Associate Professor of Sociology at UC Irvine; and Kenneth A. Dodge, William McDougall Professor of Public Policy and Professor of Psychology and Neuroscience at Duke University. All questions regarding this issue should be directed to Suzanne Nichols, Director of Publications, at journal@rsage.org and not to the email addresses of the editors of the issue.

A conference will take place at RSF in New York City on January 19, 2018. The selected contributors will gather for a one-day workshop to present draft papers (due on December 17, 2017, a month prior to the conference) and receive feedback from the other contributors and editors.

Travel costs, food, and lodging will be covered by the foundation. Papers will be circulated before the conference. After the conference, the authors will submit their final drafts on or before March 15, 2018. The papers will then be sent out to three additional scholars for peer review. Having received feedback from reviewers and the RSF board, authors will revise their papers before August 16, 2018. The full and final issue will be published in spring 2019.

Papers will be published open access on the RSF website as well as in several digital repositories, including JSTOR and UPCC/Muse.