

Population Association of America Association of Population Centers

Office of Government and Public Affairs

8630 Fenton Street, Suite 722 • Silver Spring, MD 20910
www.populationassociation.org • www.popcenters.org • 301-565-6710 x 1006



Population Association of America President

Dr. Judith A. Seltzer
University of California, Los Angeles

Vice President

Dr. Elizabeth Frankenberg
Duke University

President-elect

Dr. Amy Ong Tsui
Johns Hopkins University

Vice President-elect

Dr. Robert Hummer
University of Texas at Austin

Secretary-Treasurer

Dr. Liana C. Sayer
University of Maryland

Past President

Dr. Steven Ruggles
University of Minnesota

Dr. Lisa Berkman

Harvard University

Dr. Jason Boardman

University of Colorado at Boulder

Dr. Sarah Burgard

University of Michigan

Dr. Kathleen Cagney

University of Chicago

Dr. Marcia Carlson

University of Wisconsin, Madison

Dr. Rebecca Clark

National Institute of Child Health &
Human Development

Dr. Shelley Clark

McGill University

Dr. Parfait Eloundou-Enyegue

Cornell University

Dr. Jason Fields

US Bureau of Census

Dr. Kelly Raley

University of Texas at Austin

Dr. Michael Rendall

University of Maryland

Dr. Leah Van Wey

Brown University

Association of Population Centers

President

Dr. Lisa Berkman

Harvard University

Vice President

Dr. Pamela Smock

University of Michigan

Secretary

Dr. Eric Reither

University of Utah

Treasurer

Dr. Jennifer Van Hook

Pennsylvania State University

November 14, 2016

Commission on Evidence-Based Policymaking
Comments: Docket ID USBC-2016-0003

To whom it may concern:

On behalf of the Population Association of America (PAA) and the Association of Population Centers (APC), we are pleased to respond to the Request for Comments for the Commission on Evidence-Based Policymaking.

The PAA and APC are two affiliated organizations that together represent over 3,000 social and behavioral scientists, including demographers, economists, and sociologists, and the approximately 40 federally-supported population research centers nationwide. PAA and APC members conduct interdisciplinary scientific research and research training to further understanding about the implications of population change.

Many PAA members and affiliates of APC members are engaged in research that informs policy decision making and our understanding of the impacts of public policy on both the U.S. population and populations worldwide. Our members' research is decidedly evidence-based, making use of the best data sources, including those collected by the U.S. statistical agencies and those from other countries and international organizations. The PAA maintains a standing Committee on Population Statistics whose purpose is to continually monitor and provide input to U.S. statistical agencies on the quality of data collected by these agencies and their availability to the research community. Furthermore, many of our members have been at the forefront of developing statistical methods that are used to evaluate public policy questions in such areas as population health, fertility and mortality, and other demographic, social



and economic outcomes central to so many policy issues.

PAA and APC applaud passage of the Evidence-Based Policymaking Commission Act of 2016, which created the bipartisan Commission on Evidence-Based Policymaking and charged this Commission with:

- determining how to integrate administrative and survey data and to make those data available to facilitate research, evaluation, analysis, and continuous improvement while protecting privacy and confidentiality;
- recommending how data infrastructure, database security, and statistical protocols should be modified to best fulfill the integration and increased availability of data as described above;
- recommending how best to incorporate rigorous evaluation into program design; and,
- considering whether a federal clearinghouse should be created for government survey and administrative data.

We are pleased to continue supporting the Commission by responding to the recent request for comments published in the Federal Register on September 14. Our responses address questions (#3-7, 9, and 10) in the notice dealing with data infrastructure and access-- two issues that are central to the missions of both of our organizations and to the population sciences. Further, some of our comments are outside the scope of the questions, reflecting issues that our organizations hope the Commission will consider.

Comments: Data Infrastructure and Access

A. There are important benefits to use of administrative data, especially when linked, for conducting policy-relevant research. Administrative records have been used in a variety of research areas and provide an essential source of data for conducting important policy-relevant research. For example, such records have been used to study participation in and impacts of social programs (e.g., welfare programs, manpower training, food stamps, the earned income tax credit, etc.) on various outcomes. Often these outcomes are measured with linked administrative data, such as wage earnings (from linked unemployment insurance wage records), health conditions (from linked Medicaid records) or fertility (from linked birth certificate records). The availability of administrative records from federal, state or local sources provide a cost-effective way of supporting evaluations of these programs, regardless of whether the evaluations made



use of randomized designs for allocation of program participants to different “treatments,” or other studies that have made use of non-experimental designs.

But social program evaluation is not the only place where administrative records can and will be the primary source of data to monitor particular programs and/or evaluate particular policies or “treatments.” Furthermore, they do not only use government records. Here we reference two examples. First, biomedical research, including research that is relevant to policies affecting health-related behaviors, such as smoking bans or regulation of the nutritional content of foods, uses increasingly electronic health records (EHRs) from public and private health care systems to measure the health effects of variation in such policies. Second, administrative records from private firms that construct credit scores for use by financial institutions have been used by researchers, including the research division of the New York Federal Reserve Bank, to monitor and conduct policy-relevant research on student loan debt in the U.S. In both of these areas, administrative records support important policy-relevant research in a way that is both cost-effective and potentially more accurate than data collected via other means, e.g. surveys.

B At the same time, there are important legal and other constraints that limit the use of administrative records and the ability of researchers to link records from different administrative sources. In particular, different sources of administrative data are subject to varying and divergent laws and regulations that can inhibit their use. For example, administrative records from social programs administered at the state or local level (e.g., TANF programs) are often subject to laws and regulations that make it hard for one agency to share their records with another agency. And, as noted in the NRC report on the Reengineering of the Survey of Income and Program Participation (SIPP), existing state laws that cover the privacy and access of administrative records from TANF, Medicaid, unemployment insurance, and the workers’ compensation programs make it very difficult, if not impossible, for these programs to share their data with the Census Bureau (or other) surveys like the SIPP.¹

¹ Constance Citro and John Karl Scholz, eds., *Reengineering the Survey of Income and Program Participation*, National Research Council, 2009.



Historically, this issue has complicated the conduct of biomedical research that makes use of electronic (or non-electronic) health care records of individuals as institutional review boards (IRBs) have required studies to obtain informed consent from subjects in these studies for any follow-up use of subjects' EHRs and/or updating of these records. Recently proposed revisions to the Common Rule² will reduce and/or eliminate this re-consenting requirement for certain types of studies and types of administrative records so long as subjects are provided with a clear statement regarding potential future use of administrative records as part of their initial consent process. Many population scientists welcome this change and suggest it may represent a model for the Commission to examine as it considers how to facilitate access to records like EHRs while still providing participants with the opportunity to make informed decisions about research access to their records.

More generally, we strongly urge the Commission to investigate the various laws and regulations governing access to administrative records for research purposes. In particular, we urge the Commission to look closely at the laws affecting access to state and local government data and policies restricting record linkage across various federal agencies.

C. To facilitate the conduct of evidence-based, policy-relevant research, we urge the Commission to examine and improve access to administrative records to qualified researchers outside and inside the government. We understand and appreciate that there are important confidentiality and security concerns that necessarily limit access of researchers to various types of government-based administrative records and/or restricted-use data sources. Furthermore, we appreciate why restrictions on the access of non-governmental researchers may need to be different, and possibly more restrictive, than that applied to researchers employed by authorized government agencies. But, at times, these restrictions have made access to such data very difficult for academic and non-governmental researchers.

Over the last 20 years, U.S. statistical agencies, initially led by the U.S. Census Bureau, have made great strides in improving access to restricted-use versions of federal data sources through the Federal Statistical

² HHS–OPHS–2015–008, Federal Policy for the Protection of Human Subjects, *Federal Register*, 80(173), Sept 8, 2015, 53933-54058.

Population Association of America Association of Population Centers

Office of Government and Public Affairs

8630 Fenton Street, Suite 722 • Silver Spring, MD 20910
www.populationassociation.org • www.popcenters.org • 301-565-6710 x 1006



Research Data Centers (RDCs) program. This program now allows access to data products from 12 different federal statistical agencies for qualified governmental and non-governmental researchers in 20 different centers around the country. While some the research covered by the data agreements approved for use of these centers is often not directly related to policymaking, much of it is.

A similar effort for providing access to data from the Internal Revenue Service (IRS) under the Joint Statistical Program of the Statistics of Income (SOI) Division of the IRS has enabled qualified researchers to submit proposals for access to IRS data and to link it to various data for research purposes. This program has facilitated a number of highly visible and widely cited lines of research by Professors Raj Chetty (Stanford) and Emmanuel Saez (UC Berkeley). For example, Chetty and co-authors analyzed the association between income and the life expectancy of individuals in the U.S. since 2000 by linking IRS tax records on income with Social Security Administration death records.³ The findings of this research, especially the finding of differences by area in the associations of longevity by income, raises important questions about the sources of these disparities and how to alleviate these differences. Such research could not have been conducted without this program.

A large body of research shows that geography (e.g. neighborhoods) affects the social and economic well-being and health of individuals and families. But, state and local policymakers, researchers, and program officials often lack the data needed to measure differences in community environments, to isolate how neighborhood characteristics shape micro-level outcomes, or to test the efficacy of neighborhood-level interventions. Most survey data files lack such key contextual information, while most administrative data lack key demographic, socioeconomic, behavioral, and outcome information. While individual-level record linkage of survey and administrative data could provide such critical data for state and local-level evidence-based policymaking, most state and local researchers/program evaluators lack the resources to submit proposals and conduct these types of linkages and research within a RDC. The Commission should also encourage statistical

³ Chetty, R. et al. (2016). “The Association between Income and Life Expectancy in the United States, 2001-2014, *Journal of the American Medical Association*, 315(15): 1750-1766.

Population Association of America Association of Population Centers

Office of Government and Public Affairs

8630 Fenton Street, Suite 722 • Silver Spring, MD 20910
www.populationassociation.org • www.popcenters.org • 301-565-6710 x 1006



agencies and other researchers to create spatially-linked administrative and survey data that could be provided to state and local researchers/program evaluators outside of RDCs to increase evidence-based policymaking at the state and local level.

We urge the Commission to encourage expanding access of data and records from federal, state and local sources to qualified non-governmental and governmental researchers. This expansion should include state and local government researchers, whose access to data can provide support for accurately assessing needs, creating programs to address those needs, and delivering services in more cost-effective ways. While such efforts may include expanding the RDC and/or IRS's Joint Statistical Programs or similar programs, they should also include expanding access to spatially-linked administrative and survey data that could be provided outside of RDCs. Efforts also may include providing more funding for merit-based grants to undertake these projects, especially in light of the limited resources available to researchers in local governments and those working in non-governmental settings.

We also urge the Commission to consider recommending any necessary legal revisions that would allow federal statistical agencies to share data with researchers conducting evidence based research. For example, the Census Bureau's authorizing regulation, Title 13, does not explicitly recognize the use of sensitive data for conducting scientific research, be it policy-relevant or not, as a "benefit to the Bureau." Rather, Title 13 only supports data access to improve the quality of Census (and other) data products. A more explicit acknowledgment that qualified research projects can be conducted for scientific purposes would allow Census to approve studies using confidential data that are primarily designed to replicate existing studies and/or determine the robustness of findings from previous research. Such changes would help ensure the legitimacy of research uses of these data and give greater credibility to the findings based upon these data.

D. We encourage greater attention be given to the population representativeness of the policy-relevant research produced using administrative records and population-based surveys to better assess and characterize the population-representativeness of findings from administrative data. Many studies use administrative records to "evaluate" the impact of some particular policy or program. As we have argued above, the administrative records provide a potentially cost-effective way of conducting such evaluations—particularly when compared to the



alternative of collecting survey data that is collected from a sample representative of the population relevant for the study. However, such benefits of using administrative records in evaluative research does not mitigate the importance of assessing the sampling properties of this data source.

Consider the following example regarding the design of the Precision Medicine Initiative (PMI). One of the key components of the PMI's initial plan is to assemble a million-person sample of individuals who would provide access to their Electronic Health Records (EHRs) as a condition of the study. Access to EHRs on this large sample would provide data to study a wide range of health conditions, including conditions that are relatively rare and only affect population subgroups. One of the study's recruitment strategies was to use social media and other methods to attract participants who would grant access to their EHRs and undergo one or more physical examinations.

While the goals of the PMI are important and have the potential to provide evidence-based assessments of health conditions relevant for U.S. health policy, as population scientists, we are concerned about lack of attention to the properties of what amounts to a "volunteer" sample of people with EHRs, even if the sample includes data on one million participants. In public comments, PAA and APC raised these concerns and strongly suggested that the NIH leadership consider using existing population-based health studies to form at least part of the PMI cohort to assess the population-representativeness of the recruitment strategy based on volunteers. In developing both policies and best practices for policy-relevant research, we urge the Commission to advocate for the designs of data collection that explicitly account for the sampling properties and population-representativeness of its studies.

Lastly, we encourage the Commission to ensure that population-representative data sources collected by the Federal government continue to be viewed as an important source of data for policy-relevant research, both as a way to monitor behaviors and phenomena relevant to public policy. For example, data sources like the Current Population Survey (CPS), the American Community Survey (ACS) and the Survey of Income and Program Participation (SIPP) all play roles in the monitoring and implementing public policy in the U.S. The CPS is the population-representative data that enables the BLS to construct estimates of unemployment and labor force participation rates of the U.S. population on a monthly basis. The ACS provides data on poverty rates at the lowest

**Population Association of America
Association of Population Centers**

Office of Government and Public Affairs


8630 Fenton Street, Suite 722 • Silver Spring, MD 20910
www.populationassociation.org • www.popcenters.org • 301-565-6710 x 1006



levels of geography, such as school districts and communities, which are used to allocate funding for programs such as the USDA's National School Lunch Program and State Children's Health Insurance Program (SCHIP). The SIPP has facilitated a broad range of research on the distribution of income and participation in a range of social programs using a survey that is designed to be population representative for most states in the U.S. These surveys, and others, are important components of the U.S. data infrastructure and are needed to support evidence-based policymaking.

Thank you for considering our comments and for the important work of the Commission. Please do not hesitate to contact the Director of Government and Public Affairs, Mary Jo Hoeksema (maryjo@popassoc.org), if our organizations can be of further assistance.

Sincerely,



Judith A. Seltzer
President, Population Association of America



Lisa Berkman
President, Association of Population Centers