Access to Innovation: Utilizing the Census Bureau Data Linkage Infrastructure for Evidence-Based Research

Kate McNamara
Evidence Building Staff
Economic Reimbursable Surveys Division
U.S. Census Bureau

Beyond the Numbers Economic Data Conference
11/8/2018

United States™ Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Outline

- Introduction to the Census Bureau Evidence Building Staff

- Administrative Data Linkage Infrastructure

- Evidence-Building Projects

- Metadata and data documentation efforts

Evidence Building Staff

- Created in October 2017
- Expand the Census Bureau's Data Linkage Infrastructure
  - Acquire new data for evidence building
  - Effect policy changes to support evaluation - Title 13 Benefit Policy
  - Improve metadata quality and access
- Facilitate Government Evidence Building/Program Evaluation Research
  - Develop and document policies and processes for evidence building research at Census
  - Manage internal evidence building projects
  - Provide research support
- Streamline Project Approvals

Community Context

- Commission active from 2016-17. One year anniversary of the report September 7, 2018

- Next steps: Federal Data Strategy working groups will create a coordinated approach to federal data use

- **Recommendations:**

  - Increase use of data for policy while improving privacy

  - Creation of a National Secure Data Service to support secure access to datasets

  - Strengthen government's capacity for program evaluation

United States **Census** Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Census Bureau Mission

The Census Bureau's *mission* is to measure the nation's people and economy. We honor privacy, protect confidentiality, share our expertise globally, and conduct our work openly.

How do we realize our mission in the context of evidence-building?

- **Increase the awareness** and acceptance of administrative data in federal statistics

- **Identify, acquire, ingest, process, link, and analyze** administrative data

- **Create products** that demonstrate the value of data linkage and linked data

United States™ Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Census Collects Data From Many Sources

- Directly from households and businesses (surveys and censuses)
- Administrative Records
  - Federal, State, and Local Governments
  - Commercial Sources
- Required by law to reuse data from other agencies
  - Lower survey costs
  - Reduce respondent burden
- ALL data collections are confidential

Census Bureau Surveys

- Decennial- Population survey every 10 years

- American Community Survey- annual

- Economic Census- every 5 years

- Annual Business Survey

- And many more

Census Data Checkup: Friday, 10:00-10:45

Selected Administrative Data Sources

Federal Administrative Records	State Administrative Records	Third Party Data
• Health and Human Services • Housing and Urban Development • Social Security Administration • Bureau of Justice • U.S. Postal Service • Office of Personnel Management • Selective Service • FEMA	• Supplemental Nutrition Assistance Program • Unemployment Insurance • Women, Infants, and Children • Low Income Energy Assistance Program • Temporary Assistance for Needy Families • Child Care Subsidy • Permanent Fund	• Contact • Household • Housing

For a full inventory, see:
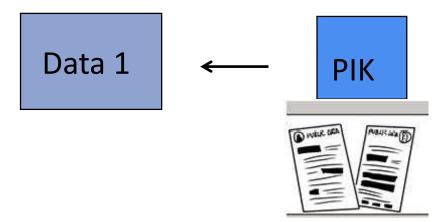www.census.gov/about/adrm/linkage/guidance.html

Key Economic/Financial Datasets

- **Business Register files**: Series of Annual files of the universe of businesses that filed with the IRS from the 1970s to today

- **The Longitudinal Business Database (LBD)**
  - Links the establishments in the BR over time
  - Allows calculation/observation of key changes

- **The Economic Census**

- **County Business Patterns**

- **Administrative data:**
  - IRS Tax Data (Challenge- barriers to access)
  - California Tax data (non-title 26)
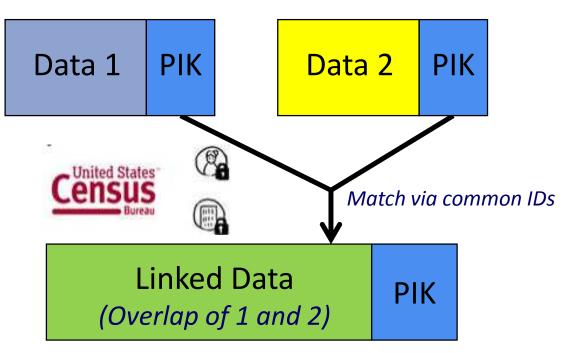 - Third party credit bureau data and realty records

United States™
Census
Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Matching person records

1. UNIQUE IDENTIFIER

*Census Bureau uses PVS to assign a unique identifier, a Protected Identification Key (**PIK**), to records in a dataset.*

2. LINKING

Using the PIK, person records are matched across datasets.
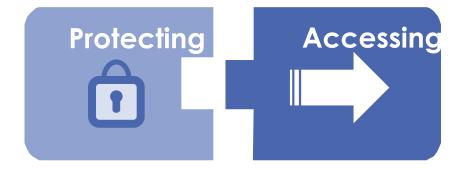

Match via common IDs

Legal framework and authority

Title 13 provides authorities for protecting and accessing high value information about the nation's population and economy.
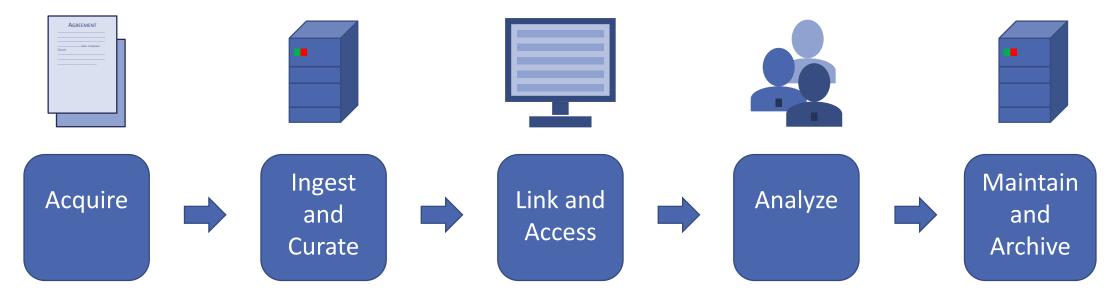
§ 6: Acquire and utilize records to the greatest extent possible

§ 8: Reimbursable studies and joint statistical projects

§ 9: Protect confidential individual and establishment data, limit access, and statistical uses

§ 23(c): Swear in researchers to assist the Census Bureau

Protecting

Accessing

Infrastructure

The Census Bureau acquires, ingests, curates, links, analyzes, and archives data. We strive to promote a sustainable and scalable model for accessing a range of high-value, sensitive, and confidential information.
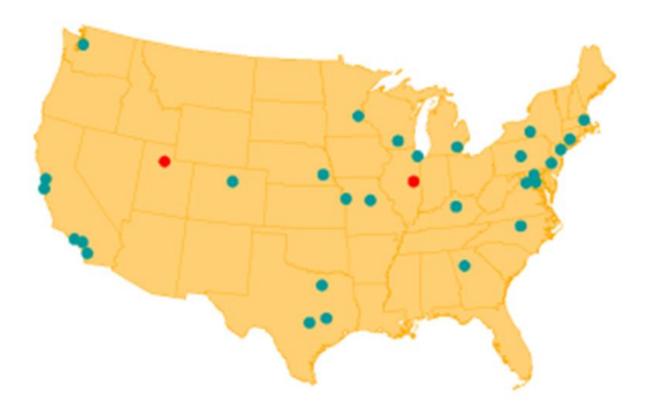
Acquire → Ingest and Curate → Link and Access → Analyze → Maintain and Archive

Secure Environment

- Data stewardship
  - Users obtain Special Sworn Status, with background check
  - Data stewardship trainings, user agreement
  - Disclosure is punishable by law

- Federal Statistical Research Data Centers (FSRDCs)
  - Physical and IT infrastructure
  - Restricted access rooms
  - Thin client monitors without data
  - NO: internet access, printers, etc.

- Disclosure review of output before release to user

Federal Statistical Research Data Centers (FSRDCs)

United States **Census** Bureau™

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Questions?

Data Linkage Infrastructure Research

- Working Papers
  https://www.census.gov/about/adrm/linkage/working-papers.html

- Projects
  https://www.census.gov/about/adrm/linkage/projects.html

- Evidence Building Projects
  https://www.census.gov/about/adrm/linkage/projects/EvidenceBuildingProjects.html

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Three Paths to Accessing the DLI

- External Project via FSRDCs

- Internal Project with Census Co-Investigator (CES, CARRA, SEHSD, etc.

- Internal Project without Census Co-Investigator (Evidence Building Staff)
  - Path for evaluation focused projects.

- Internal projects require Census Bureau review of papers and outputs

Data Discovery & Owner Approval

- **Challenges:**
  - METADATA – "What's in that file?"
  - Data owner approval

- **Remedies:**
  - ICPSR: Data Documentation Initiative
    - https://census.icpsr.umich.edu/census/
  - Streamlined release of public metadata

- **Lessons Learned:**
  - Projects that use own data move more quickly
  - Projects with phased approvals can start sooner

Legal Agreements

Challenges:

- Capacity to draft and route necessary legal agreements
- Customizing language in Agreement Templates
- Cost Recovery/ Fully Reimbursable

Lessons Learned:

- Using standard agreement language expedites the review and approval process
- Agreements with a reimbursable component clear faster
- Limited capacity for facilitating data owner approval

SSS and FSRDC Access

- **Challenges:**
  - Not everyone is eligible for Special Sworn Status (SSS)
  - Location of FSRDC vs. Location of Researchers
  - Relocation of researchers during Project Timeline
  - Seat availability
  - Communicating with research team about restricted data
  - Data storage
- **Lessons Learned:**
  - Need to continue to explore new methods for data storage and access that don't compromise data security

Current Evidence-Building Projects

- **FEMA**: Individual Assistance (IA) Administration

- **VES/AIR**: Post 9/11 GI Bill Outcomes

- **Army**: Socioeconomic Well Being of Army Personnel and Families

- **Cornell/Berkeley**: Long-run Impacts of Welfare Reform

- **RAND**: Effects of Public Need-Based Aid for College

- **CCSU/Sienna/UCONN**: Program Utilization by Formerly Criminalized Youth

- **Chicago**: Adult Labor Market Outcomes of Chicago Public School Students

- **UNC Charlotte**: Trajectories and Service Utilization of Families and Children Experiencing Homelessness

- **Nebraska/Chicago/Yale**: Does Eviction Cause Poverty?

Health at Birth and Later Life Outcomes

- L Wherry (UCLA) and S Miller (Michigan)

- Links CA Birth Records, Census, American Community Survey, SSA Administrative data, Medicare, Medicaid, TANF, WIC-CA, and IRS tax data

- Examines impact of the Comprehensive Perinatal Services Program (CPSP) and Women, Infant and Child (WIC) program on health at birth and later life outcomes

Does Eviction Cause Poverty?

- W van Dijk (University of Chicago), D Tannenbaum (University of Nebraska Lincoln, JE Humphries (Yale), N Mader (Chapin Hall at the University of Chicago

- Links Cook County court eviction cases and Chicago Homeless Management Information System to decennial Census, American Community Survey, HUD PIC and TRACS, Medicare, and Experian.

- Will evaluate the impact of eviction on employment and education outcomes

Evaluation Projects

- Moving forward, we are looking to facilitate pro
  projects
- Evaluate a federal, state, or local government p

Hosting MTO and FOS

- Host two housing randomized evaluation datasets from the Department of Housing and Urban Development (HUD)
  - **Moving to Opportunity** (MTO) data
    - Conducted in 1990s
    - Multiple follow ups, studied outcomes across social domains
  - **Family Options Study** (FOS) data
    - 3-year study conducted 2010 to 2012
- First experimental study data in the Census Bureau Data Linkage Infrastructure
- Three new projects in 2018

FEMA

- Descriptive Analysis of **Individual Assistance (IA) Program** registrants (2007 – 2017)

- To validate current data collection methods, improve predictive forecast modeling and better deliver services to populations with special needs.

- Registration and Eligibility Assessment: Pattern analysis on Individual Assistance program registrations across two main IA factors: Registrants vs. non-Registrants and, among registrants, eligible vs. ineligible for additional disaster relief programs.

- Multiple Descriptive Studies to understand the relationship between population characteristics from Census Bureau data and FEMA's disaster survivor program registration and eligibility data.

Questions?

Metadata Working Group

- Formed in February 2018 to make recommendations on metadata standards for data available in the FSRDC network

- Improve documentation for data linkage infrastructure and allow researchers to browse metadata before proposing a project

United States™
Census
Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

DDI Format

- The Data Documentation Initiative (DDI) is an international standard for describing the data produced by surveys and other observational methods in the social, behavioral, economic, and health sciences.

- Using DDI will allow us to share metadata in a standardized format

Software challenges

- Looking to acquire software to automate DDI metadata creation

- Possibilities: Colectica and Sledgehammer

- Suggestions? Let us know!

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Data Inventory

- Current inventory available as a spreadsheet on the website

- Plans to upload metadata to ICPSR

- New inventory will be available to the public with more information on datasets.

Outreach

- SCOPE Metadata Working Group

- Federal Committee on Statistical Methodology Workshops

- Administrative Data Facilities Network
    - Working group reports
    - Data intermediary project at Georgetown
    - Community of Practice Slack

More information at
www.census.gov/datalinkage

Contact us: ERD.Data-Linkage@census.gov
Kate McNamara: Kathryn.mcnamara@census.gov

Working Papers:
www.census.gov/about/adrm/linkage/
working-papers.html

References

- DLI Website
- PVS papers- 1, 2
- Federal Data Strategy
- ADRF Network