



# Federal Statistical Research Data Centers: Facilitating Research while Protecting Privacy

Joelle Abramowitz, Ph.D.  
University of Michigan

2020 Census: Citizenship, Science, Politics, and Privacy  
October 31, 2018

# Data access and protecting privacy are both important

- Federal agencies administer censuses and surveys and collect information from administrative records
- These activities produce a wealth of information that is used in a myriad of ways, from community planning to academic research
- Safeguarding the information of individuals and firms providing these data is a priority

# RDCs facilitate access to restricted data while protecting privacy

- To balance access to data and safeguarding privacy, some data are made available publicly, while others are only made available on a restricted basis
- Federal Statistical Research Data Centers (RDCs) were established to provide access to restricted data

RDCs enable **qualified researchers** with **approved projects** to access **confidential, unpublished data** from the federal statistical system

“This Research Data Center allows us to ***engage researchers*** outside of Washington in using this very important data while also ***protecting the public’s right to privacy.***”

- Former U.S. Census Bureau Director C. Louis Kincannon

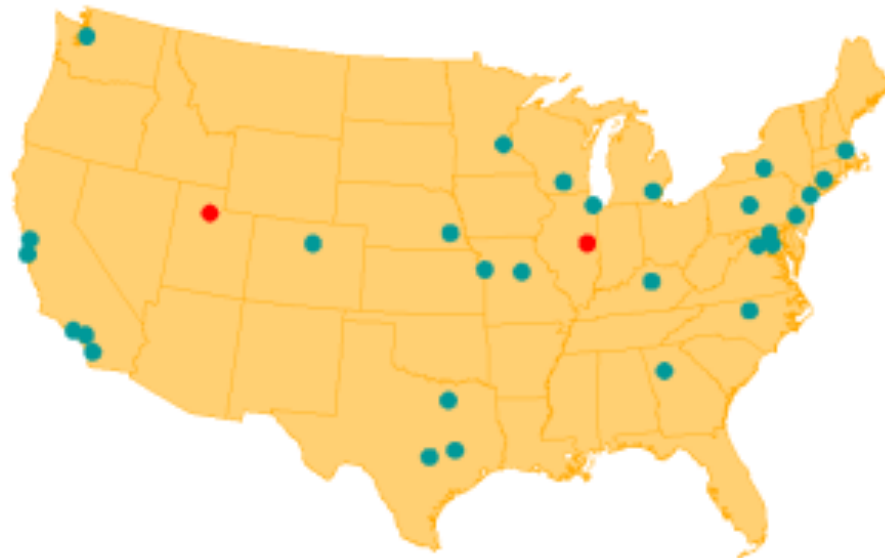
# How RDCs protect privacy

- RDCs provide access to restricted Census data and data from other federal statistical agencies
- Each RDC is a secure facility
- All RDC research output goes through a rigorous disclosure avoidance review process to ensure no confidential information is released
- Each RDC is staffed by a Census Bureau employee

# Each RDC is part of the RDC network

- The RDC network includes 29 RDCs around the country
- RDCs are joint projects of the U.S. Census Bureau and their home institutions
- The Michigan RDC is a joint project of the U.S. Census Bureau and the University of Michigan

The RDC network permits researchers to access data across RDCs and facilitates collaboration across locations



Source: U.S. Census Bureau



# Working in an RDC provides access to data not available elsewhere

- Some RDC data are not available publicly
  - Establishment-level business data
  - Linked household-firm data
- Some RDC data are available publicly, but the restricted versions provide more detailed information
  - Detailed geo-spatial variables
  - Virtually no top or bottom coding
  - Possible to link to other non-Census data

# RDCs provide Census Bureau data and, increasingly, data from other federal agencies

- Census Bureau
  - Decennial census and demographic surveys
  - Economic censuses and surveys
  - Linked business and household data
- National Center for Health Statistics and the Agency for Healthcare Research and Quality: Public health survey data
- Bureau of Labor Statistics: Labor survey data
- More agencies to join the RDC system in the future

RDCs also facilitate linkage of different data products providing unique research opportunities

- Linked business and household data
- Linked survey and administrative data

# RDCs promote broader access to data than before

- RDCs are accessible by any academic
- The application process is transparent
- RDC administrators and agency analysts assist researchers in preparing their applications
- Emphasis on timely review

# Protecting privacy: Data access requirements

- Research Proposals
  - Individuals wishing to conduct research using nonpublic data at the MRDC must submit a research proposal to the federal agency responsible for those data
- Special Sworn Status
  - Access is restricted to Census Bureau employees or researchers who have Special Sworn Status (SSS) with the Census Bureau
  - Must be either a U.S. citizen or have lived in the U.S. for three years
  - Those obtaining SSS must obtain security clearance and must sign and make a sworn statement about preserving the confidentiality of the data – for life

# Access to nonpublic data enables innovative research published in leading academic journals

Our researchers have:

- Assessed residential mobility and the geographic distribution of the healthy
- Examined the relationship between environmental emissions and health outcomes
- Quantified the impact of maternal access to the birth control pill on child poverty
- Investigated the effects of interactions with international markets on U.S. firm performance

# The RDC network is a vital resource for providing data access while protecting privacy

- Reduces the amount and scope of publicly available data to protect privacy
- Permits access to restricted data in a secure and controlled manner to facilitate important research
- RDCs are increasingly necessary as additional differential privacy measures are implemented
- Understanding the role of RDCs can be valuable in helping the public have confidence in the confidentiality of their responses to federal censuses and surveys

Thank You!

Joelle Abramowitz  
jabramow@umich.edu