

Methods and Applications of Administrative Records Research

**Sally Obenski
U.S. Census Bureau
NCVHS
September 18, 2006**

Overview

- Overview and History
- Key technical breakthroughs
- Decennial and Survey Applications
- Operational and Technical Constraints
- Conclusions

Mandate for Administrative Records Use

Title 13, Section 6:

- Use administrative records information as extensively as possible in lieu of conducting direct inquires

Census Bureau Strategic Plan:

- Reduce reporting burden and minimize cost to taxpayer by acquiring and developing high-quality data from sources maintained by other government and commercial entities

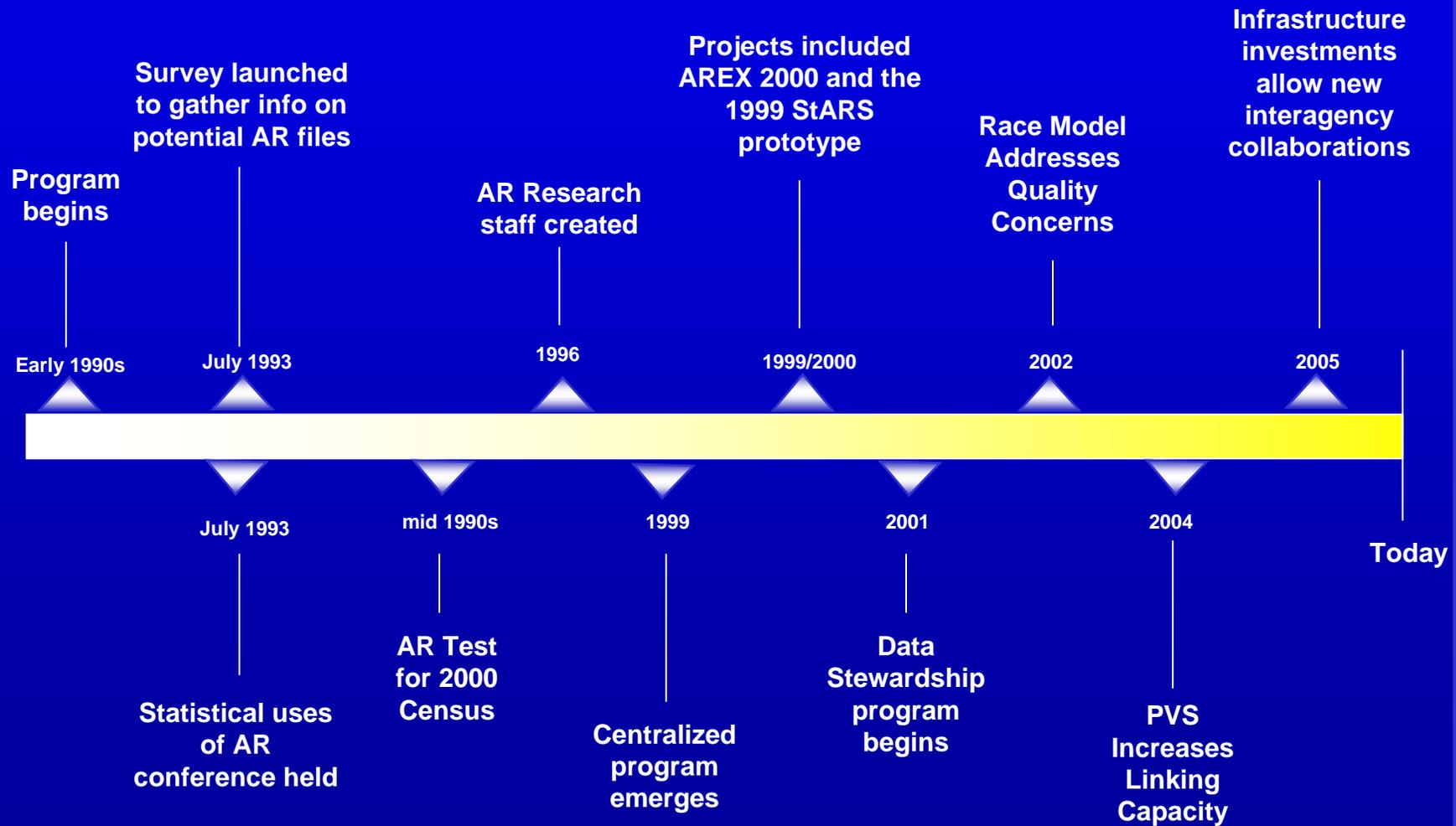
Legal Guidance and Protections

- Title 13, U.S.C., Section 6, 9, and 214
- Title 26, U.S.C., Section 6103(j)
- Privacy Act of 1974
- Paperwork Reduction Act
- Government Information Security Reform Act (GISRA)
- E-Government Act of 2002, including
 - Federal Information Security Management Act (FISMA)

Safeguarding Administrative Records

- **Consistent Application of Policies**
 - To ensure that projects have the appropriate legal authorization, comply with existing data agreements, and provide adequate controls to protect confidentiality and privacy
- **Administrative Controls**
 - Numerous levels of approval
 - Need-to-know access
 - Removal of identifiable information
 - Administrative Records Tracking System
 - Security and confidentiality training

AR Program Evolution



StARS Provides Technical Infrastructure

Person & Address Databases consist of 7 national files:

IRS 1040 HUD TRACS IHS SSS
IRS 1099 HUD PIC Medicare

CY2004 records	Persons	Addresses
Raw input	894 million	767 million
StARS	308 million	152 million

Administrative Records Experiment Validated StARS

- Local test of AR census models conducted in 5 counties
- Coverage issues similar to Census 2000
- Validated conformance of StARS to Census 2000 addresses & persons
- Improvements to StARS continue, including move to more real-time redesign (E-StARS)

NUMIDENT Provides National Reference File

- Social Security Administration (SSA) Numerical Identification (Numident) Transaction file with 803 million records
- Collapse to 431 million unique SSN records
- Usages:
 - Look-up file that provides demographic data
 - Social Security Numbers (SSNs) verification/validation

Race and Hispanic Origin Model Rectified Quality Concerns

- Initial weakness was dependence on race data from SSA's SSN transaction file
- Census 2000 records matched to SSN transaction file
- Model completed missing linkages

Person Validation System (PVS) Increases Linking Capacity

- Use master file of SSN/name/DOB as reference file
- Link addresses with SSN reference file
- Match incoming census or survey record using name, address, DOB
- Search within address first (high quality match)
- Search by name/DOB nationally if address search not successful
- Replaces SSN with unique identifier (PIK)

Implementing the ACS Provides Current Long Form Data

- Designed to ameliorate constraints of decennial long form data collection
- Provides means for timely analyses and estimates at small geographic areas
- Provides means to push models based on less granular surveys to smaller geographies

AR Integral to Census Bureau Programs

- Internal Revenue Service (IRS) 1040
 - Intercensal Estimates
 - Small Area Income and Poverty Estimates
- CMS Medicare and Medicaid
 - National Longitudinal Mortality Study
 - Small Area Income and Poverty Estimates
 - Small Area Health Insurance Estimates
- State Unemployment Insurance Files
 - Longitudinal Employer-Household Dynamics

Current Decennial Census Research

- Using AR to “assign” age, race, sex, Hispanic Origin, when a record can be matched
- Use AR to identify households with coverage problems
- Determine if commercially available & other lists can improve & help build GQ frame

Emerging Survey Improvement (1)

- Reducing ACS small area variance
 - Use AR as controls to adjust survey weights right after nonresponse adjustment
 - Preliminary research highly promising
- Obtaining characteristics on nonrespondents
 - Compared StARS persons to CPS responders to ensure consistency
 - Used StARS to obtain characteristics of nonresponding households

Emerging Survey Improvement (2)

- Reacting to disaster and other near-real time requirements
 - Katrina's effect on the federal statistical system and our lack of current response data highlighted need
 - Acquired the USPS National Change of Address File and FEMA's emergency management and flood insurance files
- Developing next generation StARS – near real-time measurements

Operational Constraints

- File Acquisition Complexities
 - Complex Memoranda of Understanding
 - State by state negotiation
 - Differences in content definition, quality, and program rules over time
- File lag time
 - MSIS (Federal Medicaid) lags by about 4 years
 - Most lag for about a year
 - Many applications require more near real-time response

Technical Constraints

- Obtaining the right data in the right format
- Varying rates of validation (e.g., Medicare 99%, Medicaid 91%)
- Coarseness of administrative data compared to nuances of surveys
- Measuring error

Overcoming the Constraints

- Resolving file acquisition issues may require OMB or Congressional assistance
- Lag time for general demographics addressed by National Change of Address file—planning move to Enhanced StARS for more near real-time response
- Standardized and centralized file acquisition
- New files in address search phase and SAS-based matcher increased validation rates
- Data Quality Standards team addressing measuring error in integrated data sets

Conclusions

- New files and innovations leading to expansion of AR uses
- New challenges continue to arise
 - Regular updating of billions of records to have a near real-time response system
 - Effectively acquiring state-based records
 - Understanding integrated data sets
- At incipience of a new generation of products and services

Contact Information

Sally M. Obenski

Assistant Division Chief for Administrative
Records Applications

Data Integration Division

U.S. Census Bureau

Washington D.C. 20233-8100

Email: Sally.M.Obenski@Census.Gov

Phone: (301) 763-4044

Cell: (301) 873-7615