

Health-Related Administrative Records Research

Ron Prevost
U.S. Census Bureau
NCVHS
September 18, 2006

U S C E N S U S B U R E A U

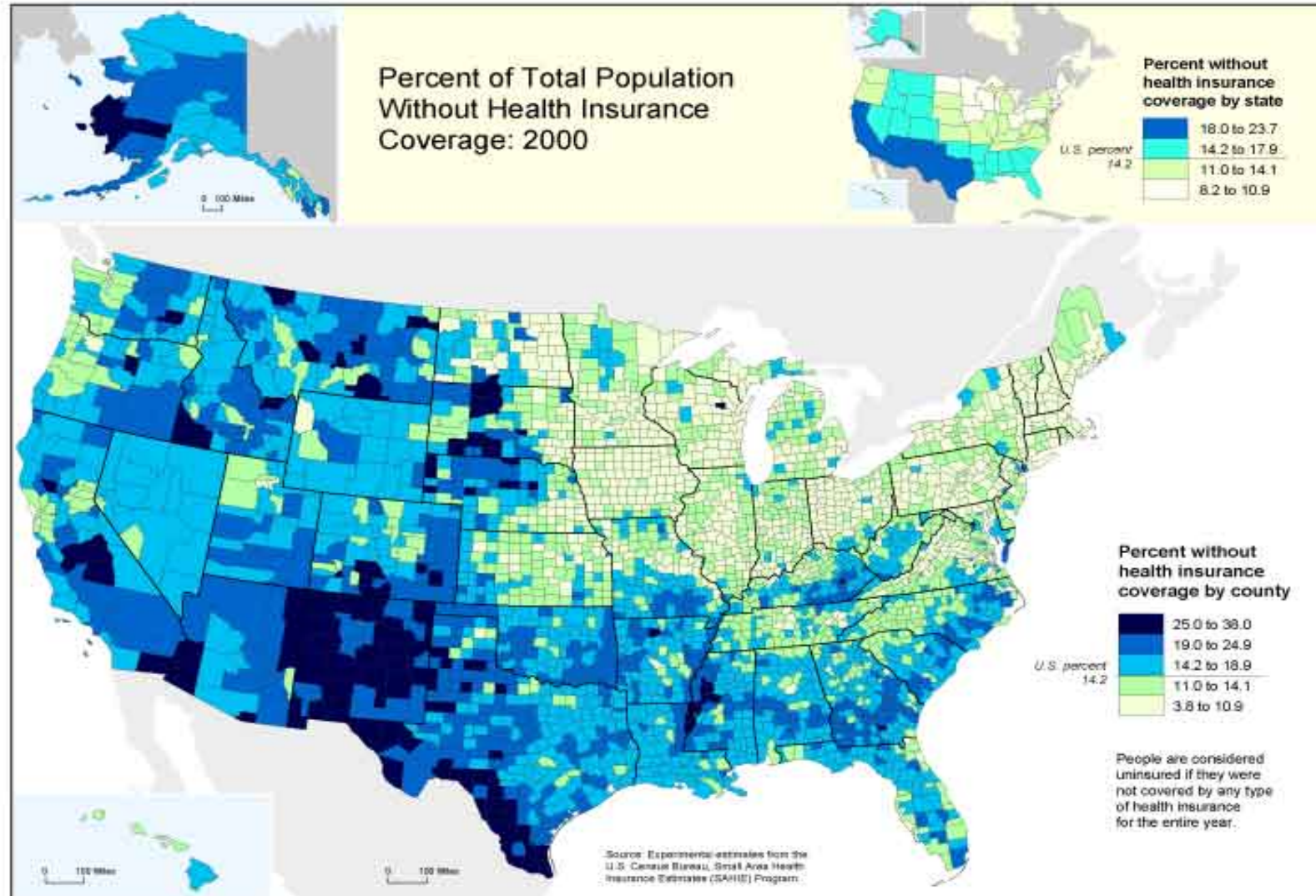
Overview

- Small Area Health Insurance Estimates
- Medicaid Undercount Study Description
- Preliminary Results
- Next Steps
- Related Research
- Benefits of Integrated Data Sets
- Policy Challenges
- Conclusions

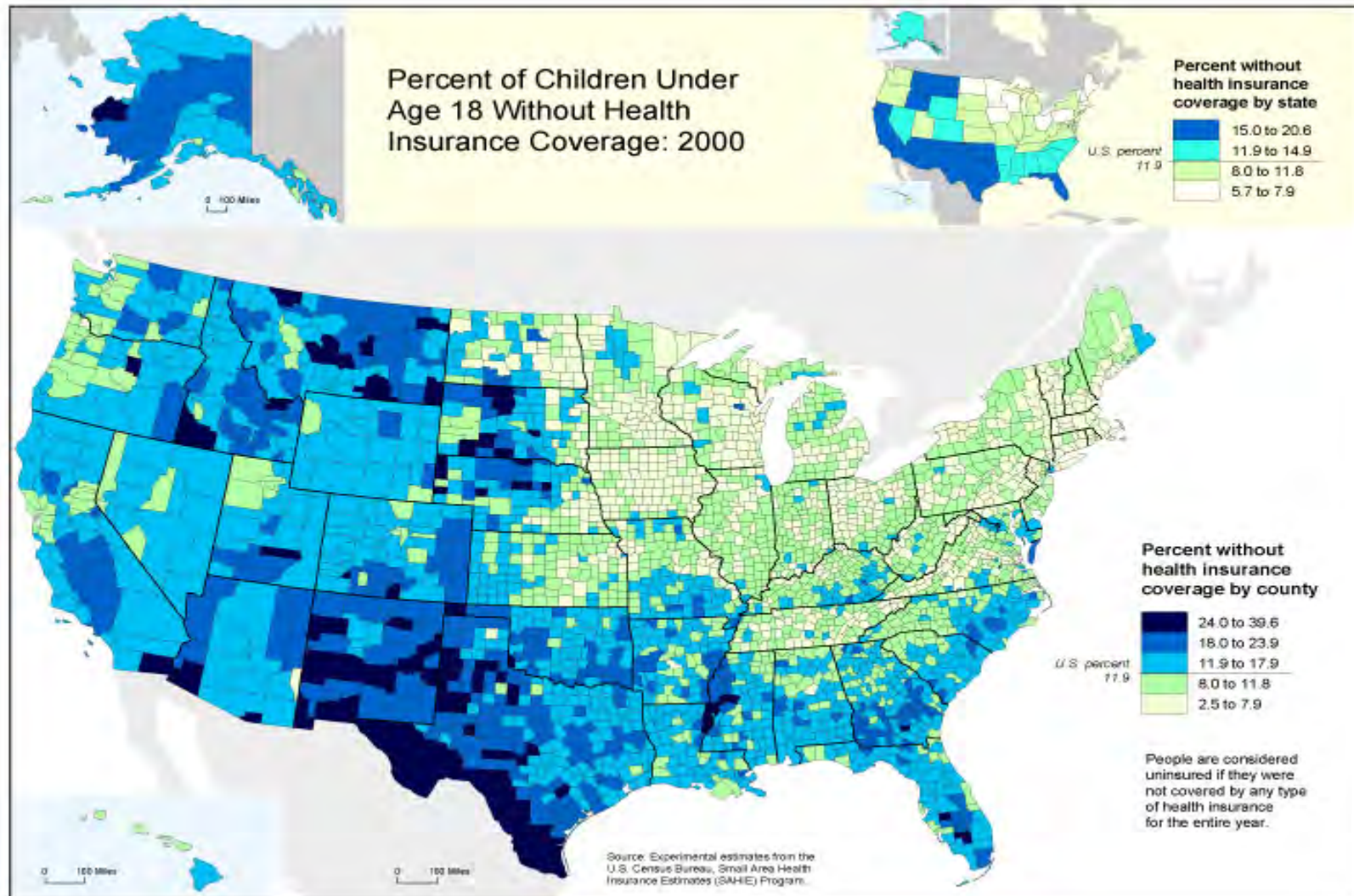
U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE) Program Overview

- Produces a consistent set of estimates of health insurance coverage for all counties
- Have published estimates for counties and states by age (under age 18, and total) with confidence intervals
 - Investigating model improvements
 - Expanding age categories

Small Area Health Insurance Measures



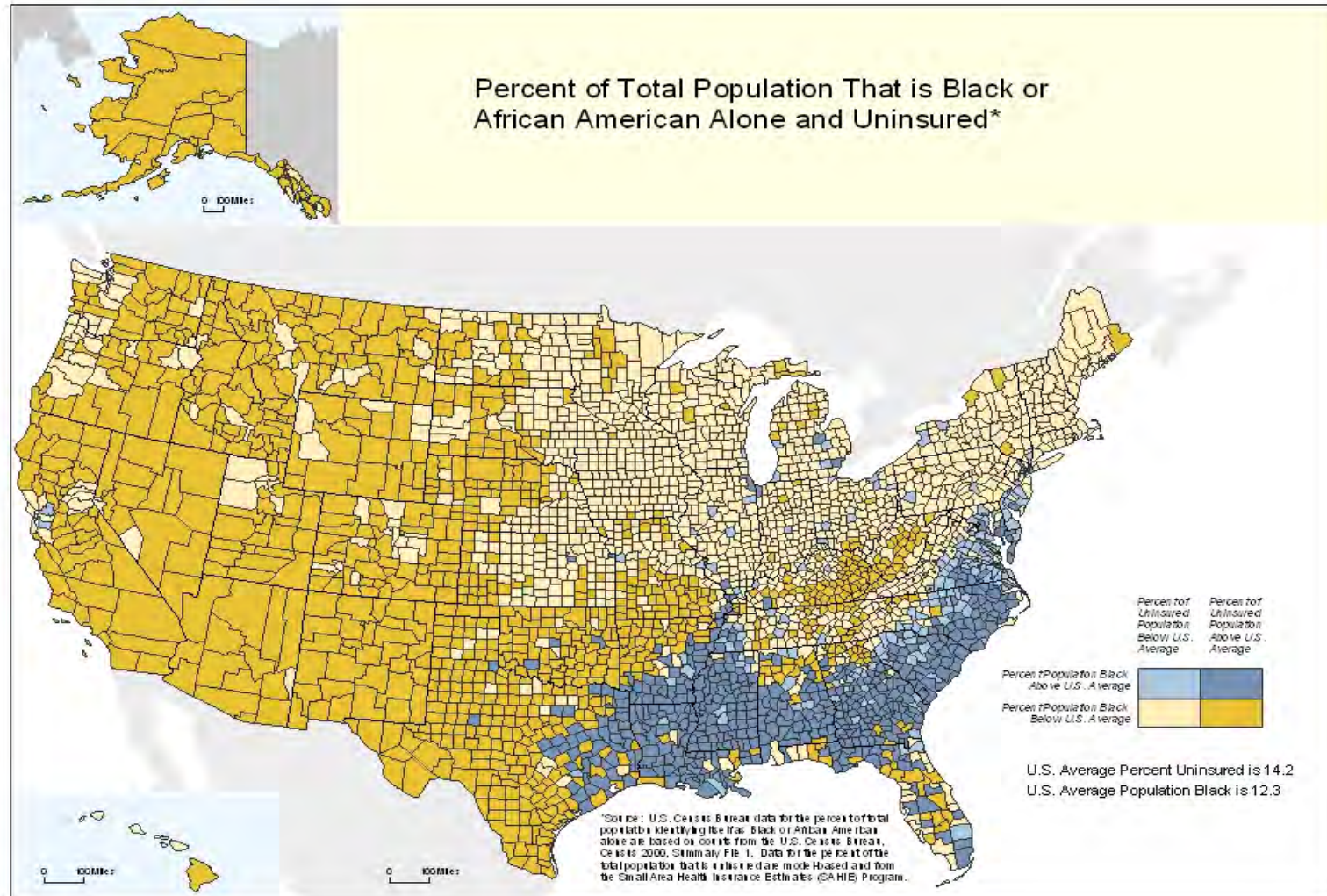
Small Area Health Insurance Measures



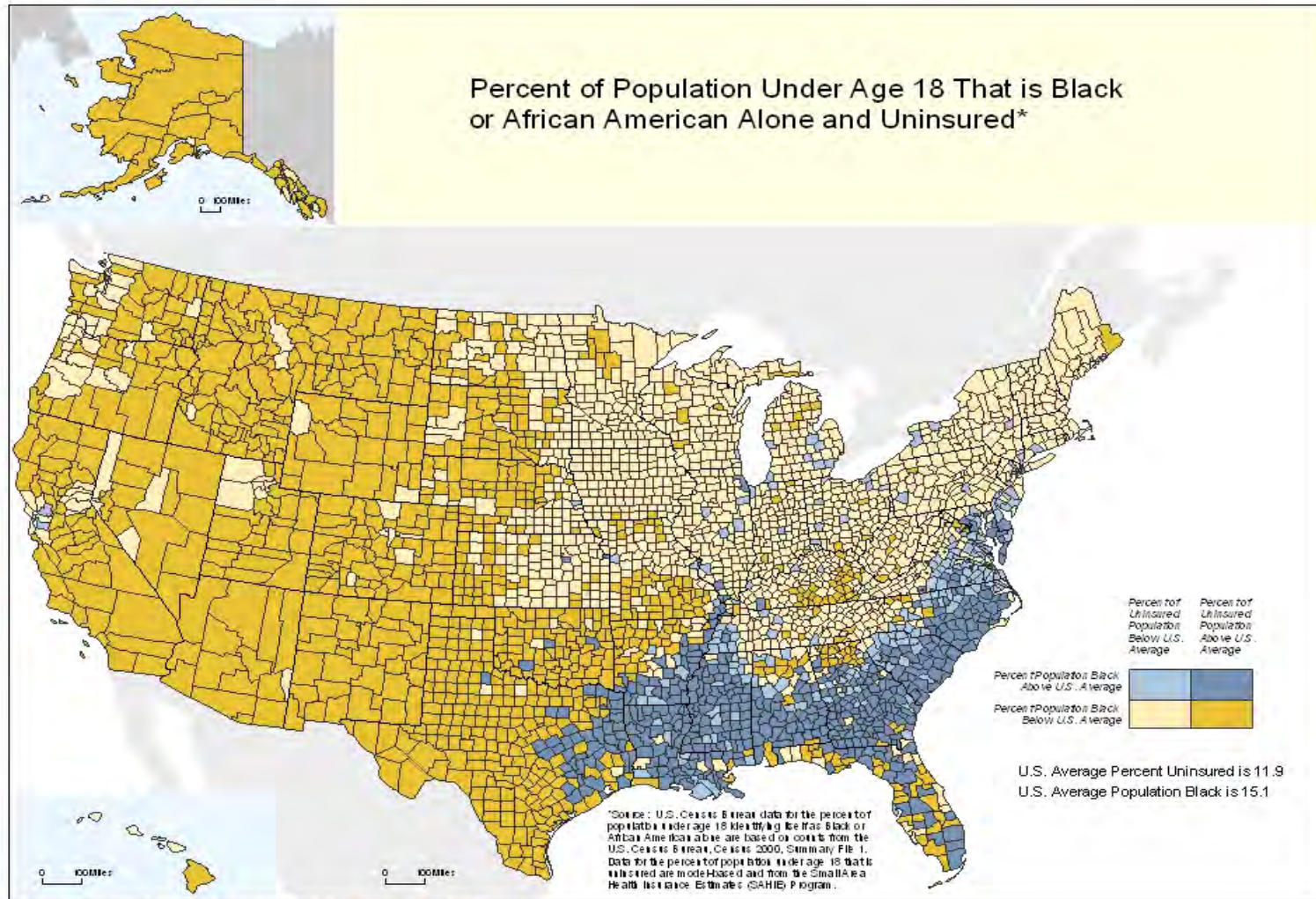
U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE) Program Overview

- Health insurance coverage estimates are created by combining survey data with population estimates and administrative records
- Race, ethnicity, age, sex and income categories are being investigated for counties and states
 - State-level estimates such as uninsured Black or African American children under age 18 \leq 200% of poverty
 - County-level estimates such as uninsured children under age 18 \leq 200% of poverty
 - This project is partially funded by the Centers for Disease Control and Prevention's (CDC's) National Breast and Cervical Cancer Early Detection Program

Small Area Health Insurance Measures



Small Area Health Insurance Measures



USCENSUSBUREAU

USCENSUSBUREAU

U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE) Program Overview

- Forthcoming health insurance coverage estimates in 2007
 - Updated county- and state-level estimates by age
 - State-level estimates by race, ethnicity, age, sex, and income categories
- Depending on future funding, the SAHIE program plans to produce county- and state-level model-based estimates as an annual series

Medicaid Undercount Project Collaborators and Co-Authors

- Project Lead: Dr. Michael Davern, State Health Access Data Assistance Center, UMN

- Collaborators:
 - US Census Bureau Collaborators:
 - Sally Obenski
 - Ron Prevost
 - Dean Michael Resnick
 - Marc Roemer

 - Office of the Assistant Secretary for Planning and Evaluation:
 - Rob Stewart
 - George Greenberg
 - Kate Bloniarz

- Coauthors:
 - Centers for Medicare and Medicaid Services
 - Dave Baugh
 - Gary Ciborowski

 - State Health Access Data Assistance Center
 - Kathleen Thiede Call
 - Gestur Davidson
 - Lynn Blewett

 - Rand Corporation
 - Jacob Klerman

What is the Medicaid Undercount?

- Survey estimates of Medicaid enrollment are well below administrative data enrollment figures
- Preliminary numbers for CY 2000:
 - Current Population Survey (CPS) estimated 25M
 - The Medicaid Statistical Information System (MSIS) estimated 38.8M
- There is a substantial undercount in the CPS relative to MSIS (in this case 64%)

Why Do We Care?

- Survey estimates are important to policy research
 - Used for policy simulations by federal and state governments
 - Only source for the number of uninsured
 - Only source of the Medicaid eligible, but uninsured population
 - Used in the SCHIP funding formula
- The undercount calls the validity of survey estimates into question
- Study is intended to understand causes

What Could Explain the Undercount?

- Universe differences between MSIS and CPS survey data
- Measurement error
- Administrative and survey data processing, editing, and imputation
- Survey sample coverage error and survey non-response bias

Issues to be Addressed

➤ **Universe Differences**

- Persons included in MSIS but not in CPS
 - Persons living in “group quarters”
 - Persons who do not have a usual residence

➤ **Persons receiving Medicaid in 2+ states**

➤ **What is “meaningful” health insurance coverage**

- Persons with restricted Medicaid benefits
- Persons with Medicaid coverage for one or few months

➤ **CPS respondent knowledge**

- Because of plan names, enrollees in Medicaid prepaid plans may not know they have Medicaid coverage

Issues to be Addressed

- **CPS responses and respondent recall**
 - Medicaid enrollees who did not use Medicaid services may not consider themselves covered by Medicaid
 - Proxy responses for other members of a household may be incorrect, especially for non-family members or multiple families
 - Persons with other insurance and/or personal liability, in addition to Medicaid, may respond incorrectly
 - Respondents may confuse Medicare and Medicaid benefits
- **Possible bias**
 - Non-response may be higher for poorer populations
 - Persons who did not receive Medicaid benefits during the month in which they were CPS respondents may not report Medicaid

Medicaid Undercount Study (1)

- **Develop Validated National CMS Enrollment File**
 - Determine coverage & validation differences between MSIS and MEDB
 - Determine characteristics of MSIS, MEDB, & Dual Eligible individuals
- **Conduct National Medicaid to CPS Person Match**
 - Determine why Medicaid and CPS differ so widely on enrollment status
 - Build suite of tables detailing explanatory factors & characteristics

Medicaid Undercount Study (2)

- **Complete State Database Work**
 - Determine effect of including State Medicaid Files on validation
 - Conduct frame coverage study using MAF/MAFARF/CPS/SS01/ and Medicaid addresses
 - Conduct state person coverage study to examine if same explanatory factors apply as in national
- **Conduct National Medicaid to NHIS Person Coverage Study**
- **Results of each phase documented in working papers**

Preliminary Explanations

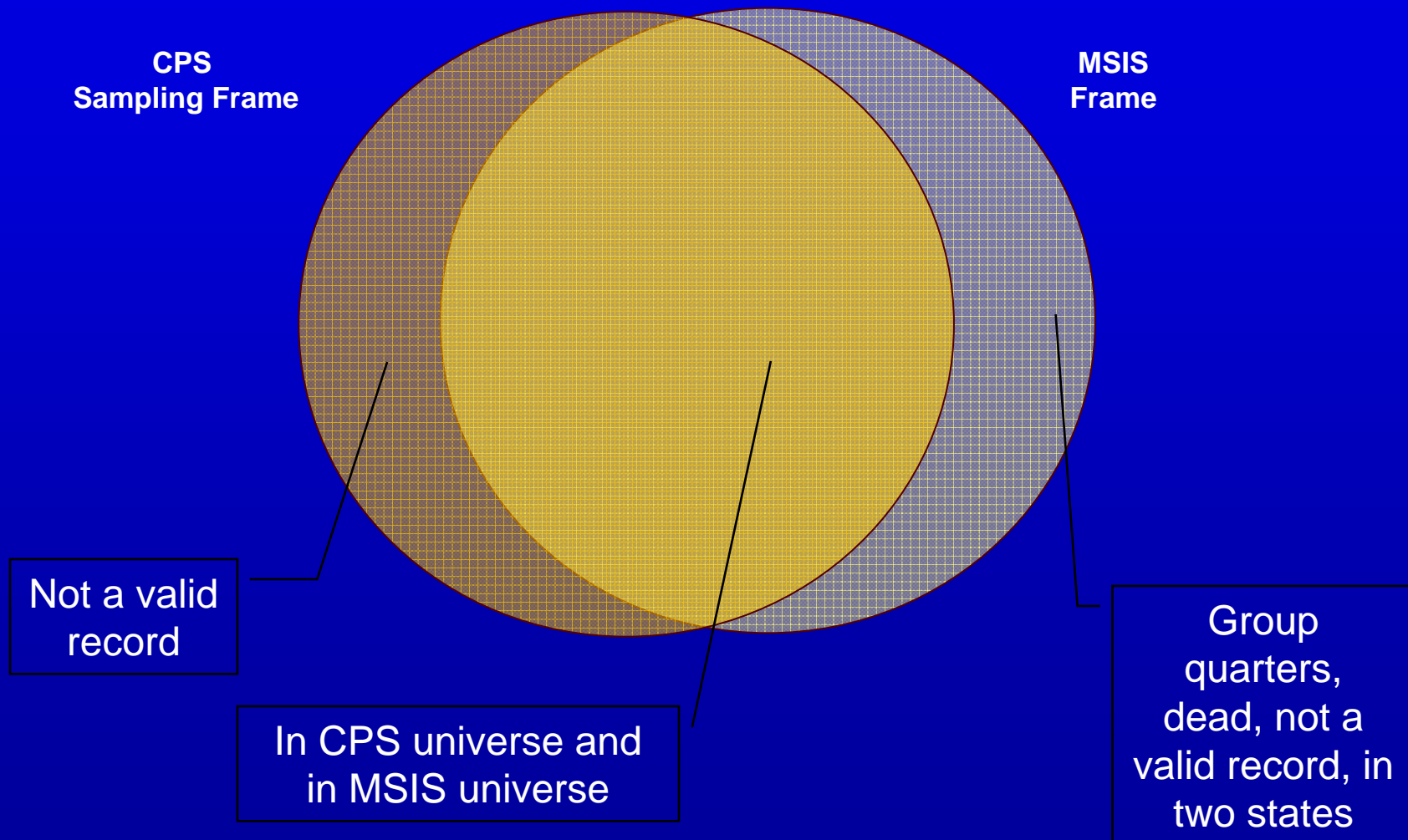
➤ Universe differences

- Enforce CPS group quarter definitions on MSIS where we have administrative data address information
- Look for duplicate persons in different states or same state

➤ Measurement error

- Link CPS respondents to MSIS data for CY 2000 to examine survey reports of enrollees
- Understand the covariates of misreporting

Building a Common Universe



MSIS Linkage to CPS

- Removed MSIS dual eligible cases defined as a “group quarter” by Census
- Ran the 2000 MSIS data through Census Bureau’s Person-ID validation system
 - A record is “valid” if in the appropriate format and demographic data is consistent
- Removed duplicate valid records
- Removed those MSIS enrollees not enrolled in “full benefits”

Matching the CPS Universe

Number of MSIS Medicaid Records in
2000:

44.3 M (total MSIS records)

- 1.5 M (records in more than one state
or group quarter)

- 4.0 M (partial Medicaid benefits)

38.8 M (the target Medicaid total)

Sample Loss in Linkage

➤ MSIS

9% of all MSIS records did not have a valid record and were not eligible to be linked to the CPS

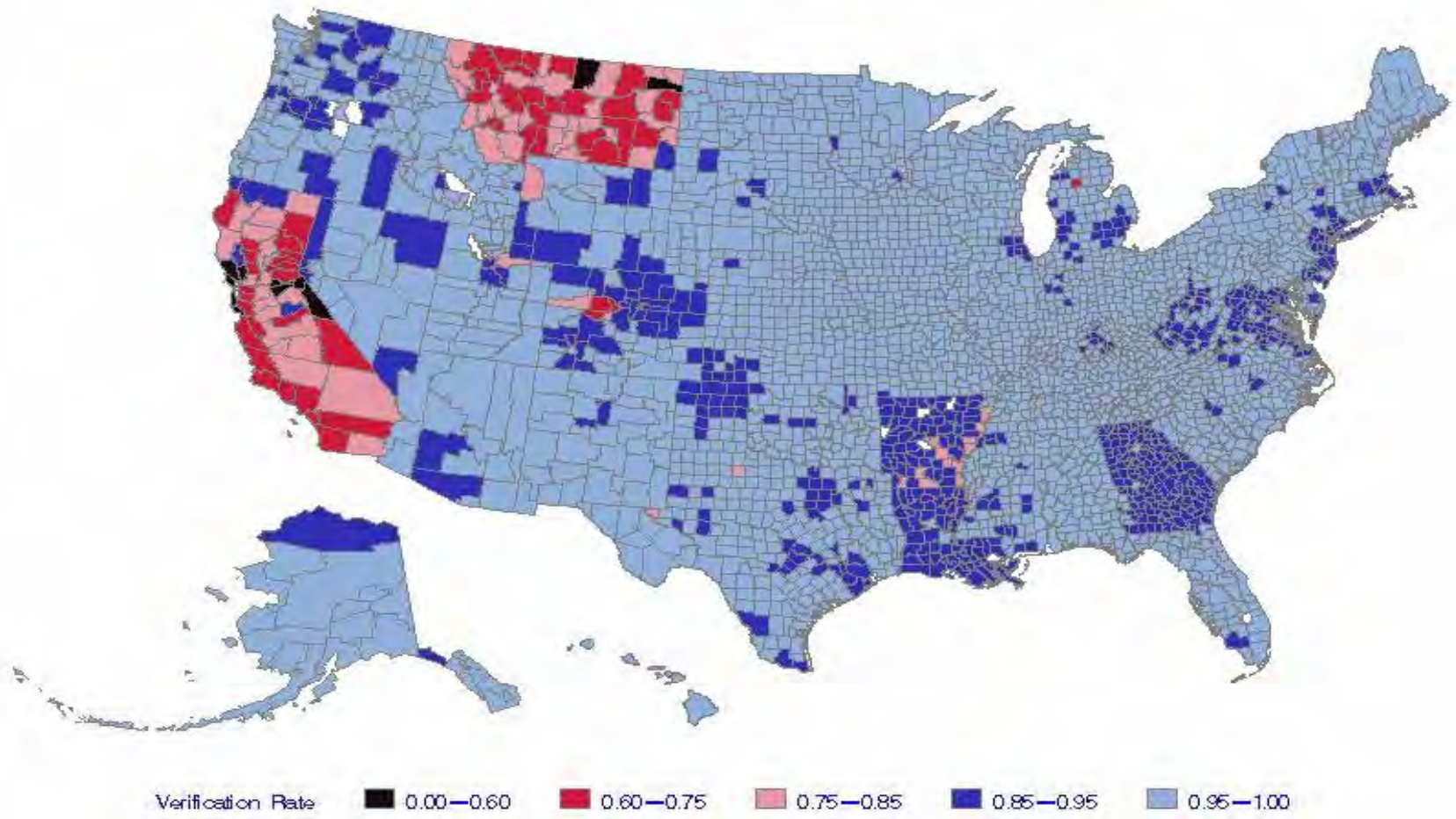
➤ CPS

6.1% (respondents' records not validated)
+ 21.5% (respondents refused to have their
_____ data linked)
27.6% (total not eligible to be linked to MSIS)

Validated Records in MSIS

SHADAC Project Phase 1 Question 1

County-Level Medicaid SSN Verification Rates for Calendar Year 2001



Matched CPS-MSIS Respondents with Reported Data Only

- 12,341 CPS person records matched into the MSIS
 - 1,906 records had imputed or edited CPS data (15.5% of total).
- Focusing on only those with explicitly reported data:
 - 60% (responded they had Medicaid)
 - 9% (responded some other type of public coverage but not Medicaid)
 - 17% (responded some type of private coverage, but not Medicaid)
 - 15% (responded they were uninsured)
 - 101% (over 100% due to rounding)

Factors Associated with Measurement Error

- Length of time enrolled in Medicaid
- Recency of enrollment in Medicaid
- Poverty status impacts Medicaid reporting but does not impact the percent reporting they are uninsured
- Adults 18-44 are less likely to report Medicaid enrollment
- Adults 18-44 more likely to report being uninsured
- Overall CPS rate of those with Medicaid reporting that they are uninsured is higher than other studies
- Overall CPS rate of those with Medicaid reporting Medicaid is lower than other studies

Work Remaining (1)

- Phase III: Measure Universe Differences:
 - Use 7 Medicaid state files with name and address information to understand the impact of MSIS non-validation (one of the states is CA)
 - Use enhanced MSIS data to further analyze the CPS sample frame coverage
- Phase IV: Assess Measurement Error:
 - Compare measurement error in the CPS to the National Health Interview Survey (NHIS) by linking the NHIS to the MSIS
 - Compare measurement error in CPS to state survey experiments

Work Remaining (2)

- Administrative and survey data processing, editing and imputation
 - Evaluate how well the CPS edits and imputations work at both the micro level and the overall macro level
 - Evaluate additional state-level Medicaid data
- Survey sample coverage error and survey nonresponse bias
 - Link the address data from the 7 states to the Census Bureau's Master Address File to determine sample coverage problems
 - Assess whether those addresses with a Medicaid enrollee are more likely to not participate in Census Bureau surveys

Preliminary Results

- We have presented preliminary results that are subject to change after further investigation
- At the moment we conclude that survey measurement error is playing the most significant role in producing the undercount
 - Some Medicaid enrollees answer that they have other types of coverage and some answer that they are uninsured
- The overall goal of the project is to improve the CPS for supporting health policy analysis
 - Especially refining estimates of the uninsured

Emerging Survey Improvements

- Use integrated data sets for insight into major discrepancies between survey and federal program data
 - Food Stamps, SSI, Medicaid enrollment all suffer from about a substantive discrepancy with current survey numbers
- Provides two views of “truth”
- Informs
 - Program administration (statistical evaluations)
 - Policy development
 - Program performance measures
- Researching integrated data for reengineering SIPP

Related Research Examples

- **Maryland Food Stamp Study**
 - **SS01 reciprocity about 50% of Maryland Food Stamp Recipients**
 - **Able to explain 85 % of the discrepancy**
 - **Misreporting -- 63 % of discrepancy**

- **Child Care Subsidy Study**
 - **Develop an eligibility model for identifying cohort in SS01 who are/should be receiving Child Care Subsidy**
 - **Researchers at RDC examining effects of CCS on employment and self support**

High-Value Data Sources for Potential Future Integration

- State food stamp and WIC files integrated with NHANES, MEPS, and CPS Food Security Supplement
- SSA's SS-5 that includes relationships not found in the NUMIDENT
- Federal Health Insurance Administrative Data Files
- State Medicaid files for inclusion in Medicaid Undercount Study

Value of Integrated Data Sets

- Provides more robust and accurate picture
- Builds on strengths of both views while controlling for their weaknesses
- Provides better statistics for input into simulations for predictions and funds distribution
- As the demand for data increases and budgets decrease data re-use may be the only cost-effective option

Policy Challenges

- Communicating the benefits vs. privacy concerns
- Need for interagency teams to ensure accurate results
- Interagency agreements and mission
- “Ownership” of the integrated data sets
- Growth of possible disclosure risks
- Need for longitudinal data bases in order to find an anonymized person at an address at a point in time

Conclusions

- Integrated data architectures are the future of American statistics
- As the demand for data increases and budgets decrease data re-use may be the only cost-effective option
- Technical and policy related challenges must be addressed
- This approach will support evidence-based public policy research and decisions.

Contact Information

Ronald C. Prevost
Assistant Division Chief for Data Management
Data Integration Division
U.S. Census Bureau
Washington D.C. 20233-8100

Email: Ronald.C.Prevost@Census.Gov

Phone: (301) 763-2264

Cell: (202) 641-2246