

Continuous Count Study

Assessing Utility and Availability of Administrative Data

Thomas Mule, J. David Brown, Marta Murray-Close,
Jerry Maples and Timothy Kennel
U.S. Census Bureau

Federal Committee on Statistical Methodology
2024 Research and Policy Conference
October 24, 2024

This research is joint work with David Brown, Adam Hall, Michael Ikeda, Timothy Kennel, Wendy Martinez, Jerry Maples, Marta Murray-Close, and the Continuous Count Study Team

DISCLAIMER: Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (CBDRB-FY24-0196).

High-Level Overview of Different Administrative Data-Based Population Estimates Projects



Demographic Analysis

- Used to estimate net coverage error in the decennial census since 1960.
- Independent of the census being evaluated.
- Developed using historical and contemporary vital records, data on international migration and Medicare enrollment records.
- At the national level, estimates are by age, sex, limited race groups and Hispanic origin.
- At the state and county levels, estimates are for young children (ages 0-4).



Population Estimates Program

- Official Census Bureau estimates.
- Annual estimates for July 1 by age, sex, race and Hispanic at the county level, and total population for subcounty places.
- Latest decennial census is the starting point for the time series.
- Administrative records and survey data are used to measure change since the date of the census in births, deaths, international and domestic migration, and housing components.



Demographic Frame

- Comprehensive database of person-level data.
- Contains demographic characteristics and addresses associated with the person.
- Derived from administrative, third-party, decennial census and survey data sources.
- Linkable to information in other enterprise frames.
- Available only to approved internal users in a secured computing environment.



Continuous Count Study

- Demonstration estimates with programmatic usages.
- Leveraging the Demographic Frame.
- Annual Administrative Population Estimates to quantify the utility and availability of administrative demographic data in years between censuses.
- Focuses on tract by demographic domains.
- Produces alternative estimates to assess the coverage of the Annual Administrative Population Estimates.

Components of the Continuous Count Study

1. Annual Administrative Record Population Estimates (AAPE)
 - Quantify availability and utility of administrative data for population estimates and 2030 Census use cases
 - Use administrative data to produce full population list each year
 - Begin with extract from Demographic Frame and add enhancements

Components of the Continuous Count Study

1. Annual Administrative Record Population Estimates (AAPE)

- Quantify availability and utility of administrative data for population estimates and 2030 Census use cases
- Use administrative data to produce full population list each year
- Begin with extract from Demographic Frame and add **enhancements**

A key enhancement: Include people whose administrative records lack geocoded tracts. Impute their tracts of residence on the reference date.

Components of the Continuous Count Study

1. Annual Administrative Record Population Estimates (AAPE)

- Quantify availability and utility of administrative data for population estimates and 2030 Census use cases
- Use administrative data to produce full population list each year
- Begin with extract from Demographic Frame and add enhancements
- Different from PEP: not postcensal estimates based on 2020 Census data
- Produce estimates from full population list and from partial lists excluding people with incomplete geographic data

Components of the Continuous Count Study

2. Comparison estimates from alternative methodologies

- Assess coverage and quality of AAPE estimates
- Example: Calibrated Administrative Record Dual System Estimates (Calibrated Adrec DSE)
 - Produce base-year estimates from 2020 Census and 2020 Demographic Frame extract
 - Move results forward using estimated tract- and county-level growth rates from Demographic Frame extracts and PEP

Domains for initial results

Reference dates

- April 1, 2020
- July 1, 2021

Geography

- Nation
- State
- Census tract

Age

(2 categories)

- 0-17
- 18+

Race and Hispanic origin

(8 categories)

- Hispanic
- Non-Hispanic White alone
- Non-Hispanic Black alone
- Non-Hispanic American Indian or Alaska Native (AIAN) alone
- Non-Hispanic Asian alone
- Non-Hispanic Native Hawaiian or Pacific Islander (NHPI) alone
- Non-Hispanic some other race alone
- Non-Hispanic multiracial

National estimates

	2020 Census (2020) / PEP (2021)	AAPE		Calibrated Adrec DSE
		Excluding imputed tracts	Including imputed tracts	
April 1, 2020				
Total population	331,449,281	325,300,000	334,900,000	332,500,000
Percent difference		-1.9	1.0	0.3
July 1, 2021				
Total population	332,031,554	324,500,000	334,900,000	333,000,000
Percent difference		-2.3	0.9	0.3

Notes: A positive percent difference means the CCS estimate exceeds the official estimate.

Source: 2020 Census Demographic and Housing Characteristics File (DHC), Population Estimates Program (PEP) vintage-2022 estimates for July 1, 2021, and Continuous Count Study.

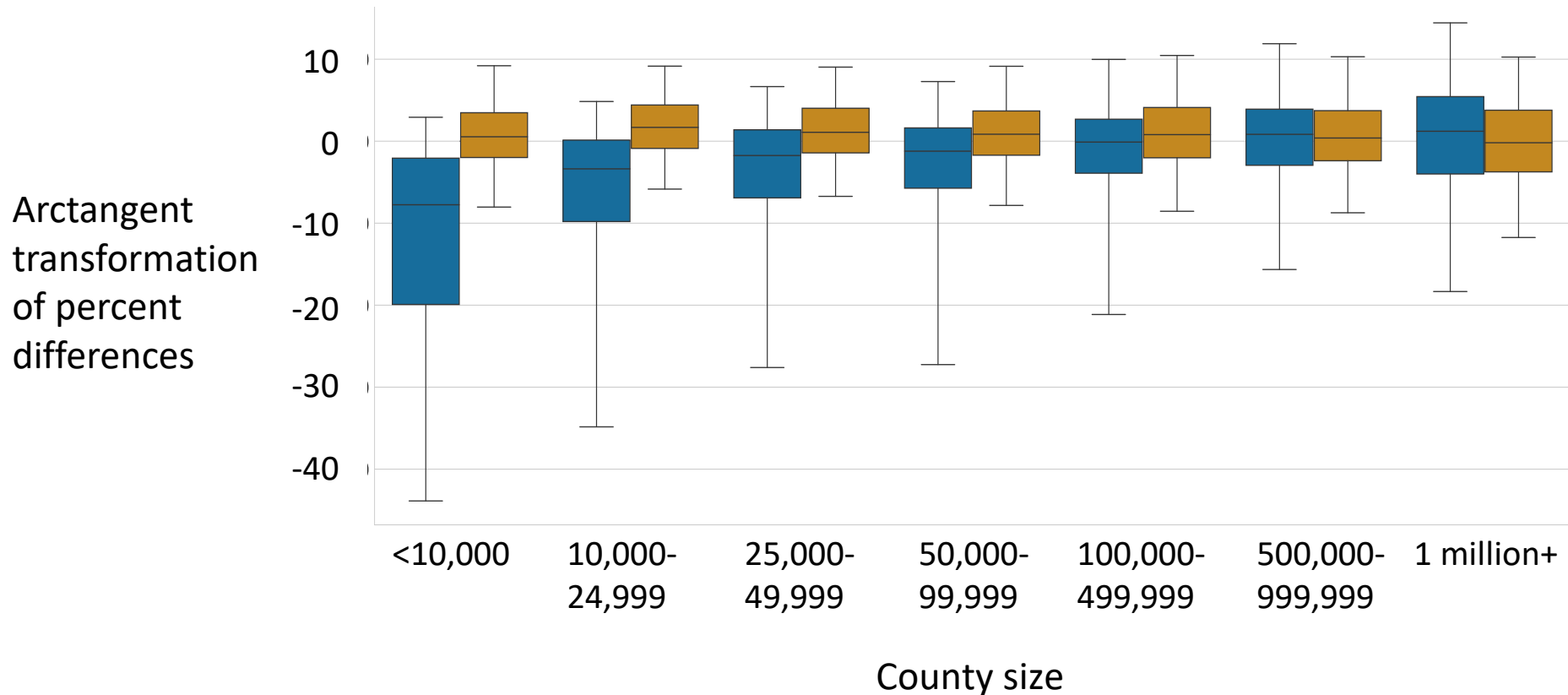
Tract-level estimates

Does the availability of administrative records for a tract depend on the size of the county where the tract is located?

We group counties by size as follows.

< 10,000	100,000-499,999
10,000-24,999	500,000-999,999
25,000-49,999	1 million+
50,000-99,999	

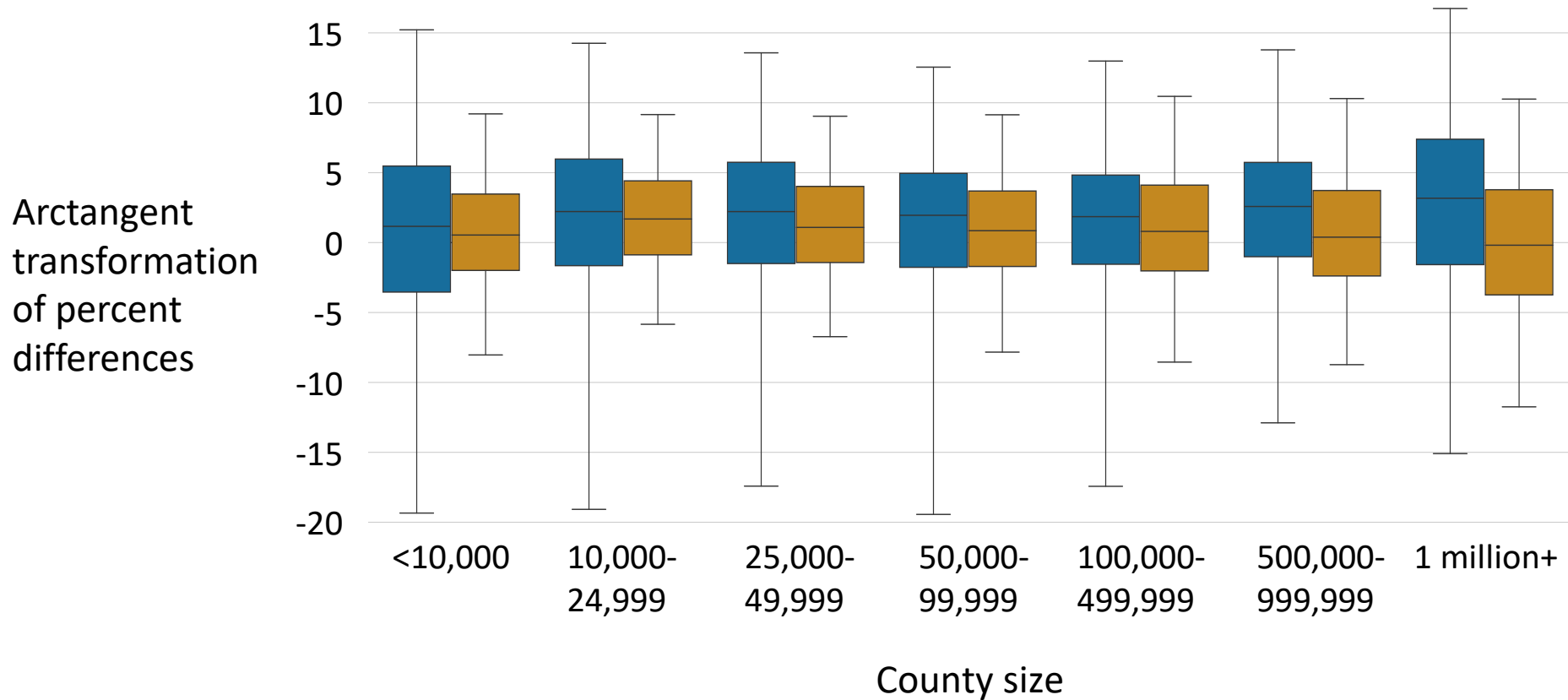
Distributions of tract-level percent difference with 2020 Census, **AAPE excluding imputed tracts** versus **Calibrated Adrec DSE**, by county size



Notes: Boxplot whiskers are truncated at 5 and 95 percent.

Source: 2020 Census Demographic and Housing Characteristics File (DHC) and Continuous Count Study.

Distributions of tract-level percent difference with 2020 Census, **AAPE including imputed tracts** versus **Calibrated Adrec DSE**, by county size



Notes: Boxplot whiskers are truncated at 5 and 95 percent.

Source: 2020 Census Demographic and Housing Characteristics File (DHC) and Continuous Count Study.

Usage of the
CCS Estimates
for the
2026 Census Test Site
Selection

Site Selection Criteria

Geography

- Locations across 2 Census Bureau Regions
- Rural area with:
 - High percentages of historically undercounted populations and
 - Historically lower self-response rates
- Area with a colonia
- Tribal area
- Area with nonmailable addresses

Living Quarters Characteristics

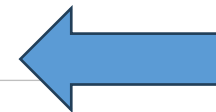
- Seasonally vacant housing
- Complex households
- High instances of new construction
- Multiunit structures and/or hidden housing units
- Wide range of group quarters types

Staffing

- Sites where recruiting field staff has been historically challenging
- Sites with different levels of unemployment rates

Technical Characteristics

- Little to no cell phone or internet service
- Comprehensive and minimal administrative data



Colorado Springs, Colorado

(Selected areas of the metro area)

Key selection criteria

- ✓ Multiunit structures and hidden housing units
- ✓ Complex housing units
- ✓ New construction
- ✓ All group quarters types, including multiple universities, correctional facilities, and military installations.
- ✓ Comprehensive and minimal administrative record data

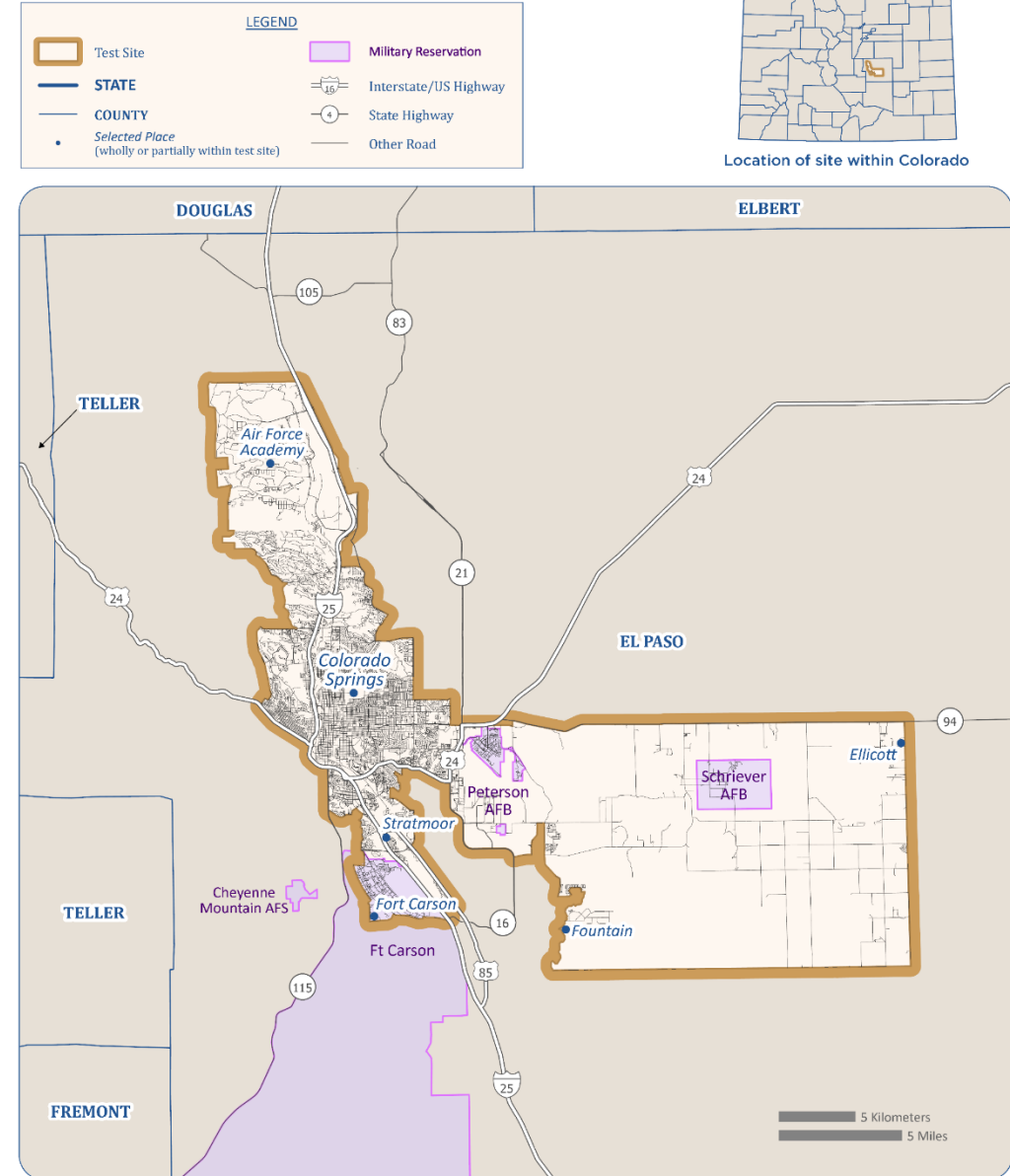
Optimal for testing

Automated group quarters procedures with correctional facilities, large universities, military installations and other group quarters types

New In-Field Enumeration procedures for complex housing units, multi-unit structures, and/or hidden housing units

Procedures for conducting early In-Field Enumeration in areas surrounding universities

Contact strategies with housing units in areas with comprehensive and minimal administrative record data



Summary of Continuous Count Study

- Produced demonstration administrative record-based population estimates for nation, states, and tracts
- Developed an estimation approach to assess estimates, including for tracts
- Produced estimates for a July 1, 2022, reference date
- Conducting ongoing research to develop the methods to produce these estimates
- Showing how results like these can help support 2030 Census planning by providing information on an ongoing basis about the availability and fitness-for-use of administrative data
- More information is available at the Continuous Count Study website

<https://www.census.gov/about/adrm/linkage/projects/continuous-count-study.html>

Thank you!

Thomas Mule
Vincent.t.mule.jr@census.gov