WHAT’S THE PRICE OF PRIVACY?
2020 CENSUS DISCLOSURE AVOIDANCE SYSTEM IMPACTS ON TRIBAL NATION CENSUS DATA

Yvette Roubideaux MD MPH
Vice President for Research and Director

Gwynne Evans-Lomayesva MPP
Researcher

Policy Research Center
National Congress of American Indians

May 18, 2021
2020 Census

• Decennial Census

• Once every 10 years – goal is a complete count
  • ACS – sample every year – estimates – social and economic characteristics

• 2020 Census questions
  • Demographic – age, sex, race/ethnicity
  • Housing data
2020 Census – AI/AN data uses

• Redistricting and representation

• Federal funding formulas and decisions

• Local tribal governance

• AI/AN research and surveillance
2020 Census Self-Response Rate

Figure 5: 2020 Census Self-Response Rates for Tribal Lands by NCAI Region
2010 final response vs. 2020 response to date

- Alaska: 31.7%
- Eastern Oklahoma: 49.4%
- Great Plains: 33.2%
- Midwest: 49.5%
- Navajo: 22.6%
- Northeast: 31.0%
- Northwest: 58.5%
- Pacific: 38.1%
- Rocky Mountain: 35.1%
- Southeast: 34.9%
- Southern Plains: 53.0%
- Southwest: 45.7%
- Western: 38.2%
2020 Census process

Census Count → Census Data Processing → Census Privacy methods → Census Data Products
2020 Census – Data privacy vs. accuracy

• US Census Bureau must protect privacy of individual census data
  • Required by Title 13
  • New issue – third party methods – potential for reconstruction/re-identification

• Disclosure Avoidance System (DAS)
  • Methods applied to the census data to keep individual data private
  • 2010 – some privacy protections applied, still published data products
  • 2020 – NEW methods - Differential Privacy/Post Processing – increased privacy protections
    • Problem: more privacy, less accuracy of the data, poorer quality data, some data not available

• Impact of more privacy protection
  • American Indian/Alaska Native data/data on tribal nations is especially impacted
  • Small, rural, remote populations also impacted
2020 Decennial Census, DAS, and Products

Census form data → Decennial Response File (Raw Data) → Census Unedited File → Census Edited File

Disclosure Avoidance System

Apportionment File to President by 4/28/2021

Group 2 Data Products, now Detailed Demographic Data: AI/AN Summary File

Differential Privacy Added Noisy Data Infused → Compiled Dataset Differentially Private Data → Post-Processing of Data

Differential Privacy Added Noisy Data Infused

Group 1 Data Products:
PL94-171 Redistricting File
Demographic and Housing File TBD

Detailed Microdata File

To States by 9/30/2021

Image Adapted From: Boyd, Danah. (2019). Differential Privacy in the 2020 Decennial Census and the Implications for Available Data Products. Adaptations based on meetings and presentations by the U.S. Census Bureau
Impact of DAS Test Four on AI/AN tribal lands (November 2020 Demo Data)


## Decennial Census and DAS Timelines

<table>
<thead>
<tr>
<th>Fall 2020</th>
<th>2020 Census Timelines</th>
<th>2020 DAS Timelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2020</td>
<td>Census Enumeration Ended 10/15/20</td>
<td>DSEP DAS Decision 11/24/2020</td>
</tr>
<tr>
<td>April 28, 2021</td>
<td>Apportionment Counts to President</td>
<td></td>
</tr>
<tr>
<td>April 30, 2021</td>
<td>Release of demonstration product 5 and 6</td>
<td>DSEP Decision Privacy Loss Budget Apply DAS to Census Data (?June)</td>
</tr>
<tr>
<td>September 30, 2021</td>
<td>PL94-171 Redistricting File to States (in August, the legacy format will be released)</td>
<td>Privacy protected data used to create all the census products</td>
</tr>
<tr>
<td>???</td>
<td>Other Group 1 Products; Group 2 Products</td>
<td></td>
</tr>
</tbody>
</table>
Census Data Privacy vs. Quality (Accuracy)

• Any loss of counts or accuracy of 2020 Census data for tribal nations is unacceptable
• Review of 2020 April Demonstration Products is the last chance for input

• Tribal consultation webinar – May 19, 2021
• Opportunity to submit responses to questions posted in the email by May 28, 2021
• Submit to: OCIA.TAO@census.gov

• Suggestions:
  • Describe how you use census data and why that use is critical for tribal governance
  • Describe why accurate census data is critical for your tribal nation – and how accurate it needs to be
  • Describe any examples of inaccurate census data in funding formulas and the impact on your tribal nation
  • Other suggestions once you see our analysis of the latest data
  • Consider the price of privacy vs. accuracy for your Tribal Nation
April 2020 Demonstration Products

• Last two data releases before final decisions by Data Stewardship Executive Policy (DSEP) Committee

• Demonstration Product 6 (PLB 4.5)
  • just changes in algorithm and processing of the data from the November release
  • Changes in how geographic spine is involved from the November release
  • Still high level of privacy, lower accuracy level of PLB 4.5 in prior releases

• Demonstration Product 5 (PLB 12.2)
  • All the changes above + increase in PLB to 12.2 – more accuracy in data

• Error Metrics also released
  • To compare changes in the level of inaccurate data in each demonstration product
  • Metrics published for AI/AN Tribal Areas/Geographies/Lands
Total Population: Percent of AI/AN Geographies where Absolute Percent Difference in PPMF exceeds 5% from 2010 public data

<table>
<thead>
<tr>
<th>Federal AI Reservations/ORTL</th>
<th>ANVSA</th>
<th>OTSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>50</td>
<td>17</td>
</tr>
<tr>
<td>32</td>
<td>38</td>
<td>17</td>
</tr>
<tr>
<td>22</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

Legend:
- Demo 4
- Demo PLB 4.5
- Demo PLB 12.2
Percent of AIAN Geographies where AIAN alone population in PPMF is less than 2010 public data

Federal AI Reservations/ORTL

ANVSA Geographies

OTSA

Demo 4  Demo PLB 4.5  Demo PLB 12.2
Risk to Small, Rural, and Remote Populations

Figure 2. Tribal Lands by Population Size

<table>
<thead>
<tr>
<th>Population Size</th>
<th>Total Number of AI/AN Tribal Lands</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 500</td>
<td>353</td>
</tr>
<tr>
<td>500 – 999</td>
<td>83</td>
</tr>
<tr>
<td>1,000 – 2,499</td>
<td>48</td>
</tr>
<tr>
<td>2,500 – 4,999</td>
<td>45</td>
</tr>
<tr>
<td>5,000 – 9,999</td>
<td>36</td>
</tr>
<tr>
<td>10,000 – 24,999</td>
<td>27</td>
</tr>
<tr>
<td>25,000 – 49,999</td>
<td>10</td>
</tr>
<tr>
<td>50,000 – 99,999</td>
<td>5</td>
</tr>
<tr>
<td>100,000+</td>
<td>10</td>
</tr>
</tbody>
</table>
AI/AN Tribal Lands with Zero Counts are Unacceptable

Figure 3. Complete Population Loss in AI/AN Areas that had a Population above Zero in the 2010 Census Data (Summary File 1) – April 2020 Demonstration Products 6 (PLB 4.5) and 5 (12.2)

<table>
<thead>
<tr>
<th>Population</th>
<th>Number of AI/AN Tribal Lands with 100% Population Loss (PLB 4.5)</th>
<th>Number of AI/AN Tribal Lands with 100% Population Loss (PLB 12.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Population</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>AI/AN Alone Population</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>AI/AN Alone and in Combination</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>
## AI/AN Tribal Lands – Population Loss of More than 5 or 10% with the Privacy Methods

Figure 4. Number of Census AI/AN Tribal Lands with Percent Population Losses in the April 2021 Demonstration Products in Demonstration Product 5 (PLB 12.2) and 6 (PLB 4.5)

<table>
<thead>
<tr>
<th>Percent Population Loss</th>
<th>Total Population</th>
<th>AI/AN Alone Population</th>
<th>AI/AN Alone and In Combination Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLB 4.5</td>
<td>PLB 12.2</td>
<td>PLB 4.5</td>
</tr>
<tr>
<td>100%</td>
<td>3</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>50 – 99.99%</td>
<td>3</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>25 – 49.99%</td>
<td>5</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>10 – 24.99%</td>
<td>30</td>
<td>10</td>
<td>64</td>
</tr>
<tr>
<td>5 – 9.99%</td>
<td>38</td>
<td>23</td>
<td>42</td>
</tr>
<tr>
<td>2 – 4.99%</td>
<td>67</td>
<td>54</td>
<td>75</td>
</tr>
<tr>
<td>&gt;0 – 1.99%</td>
<td>140</td>
<td>170</td>
<td>91</td>
</tr>
<tr>
<td>≥10%</td>
<td>41</td>
<td>18</td>
<td>109</td>
</tr>
<tr>
<td>≥5%</td>
<td>79</td>
<td>41</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Range of Gains and Losses in Counts

Figure 5. Range of Total Counts Gained or Lost in Census Data and April Demonstration Products (PLB 4.5 and 12.2) – Minimum and Maximum Values for Population Count Changes with Privacy Protections

<table>
<thead>
<tr>
<th></th>
<th>Total Population for AI/AN Lands</th>
<th>AI/AN Alone Population for AI/AN Tribal Lands</th>
<th>AI/AN Alone and in Combination Population for AI/AN Tribal Lands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLB 4.5</td>
<td>PLB 12.2</td>
<td>PLB 4.5</td>
</tr>
<tr>
<td>Largest Count Lost (Min)</td>
<td>-1107</td>
<td>-242</td>
<td>-198</td>
</tr>
<tr>
<td>Largest Counts Gained (Max)</td>
<td>1047</td>
<td>254</td>
<td>229</td>
</tr>
</tbody>
</table>
Overall Losses and Gains in Counts for Tribal Lands

Figure 6. Total AI/AN Losses and Gains in Counts – April Demonstration Product 6 (PLB 4.5) and 5 (12.2)

<table>
<thead>
<tr>
<th>Losses and Gains for AI/AN Tribal Lands</th>
<th>Total Population</th>
<th>AI/AN Alone Population</th>
<th>AI/AN Alone and in Combination Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLB 4.5</td>
<td>PLB 12.2</td>
<td>PLB 4.5</td>
<td>PLB 12.2</td>
</tr>
<tr>
<td>Overall Gain/Loss in Counts in Tribal Lands</td>
<td>-2,444</td>
<td>-1,095</td>
<td>-1,525</td>
</tr>
</tbody>
</table>
The Potential Gains and Losses are Random
U.S. Census Bureau Tribal Consultation

• Wednesday, May 19, 2021 – Webinar
• Friday, May 28, 2021 – Comments due
• Initial thoughts (NCAI will publish a letter soon)
  • AI/AN Tribal Nations are sovereign; federal trust responsibility to listen to input and ensure they are counted in the 2020 Census; government to government solutions
  • Each Tribal Nation needs to consider - What is the Price of Privacy that we are willing to bear for potential losses in census data counts?
  • Is 5 or 10% error ok, or too much?
  • No Tribal Nations should have zero counts if they had counts in the 2020 Census
  • Remember priority uses for Census data – how accurate does it need to be?
  • Should all AI/AN Tribal Lands be held invariant i.e. the actual counts?
  • More time needed after the final decision for review by Tribal Nations
Questions?
Resources

- NCAI Policy Research Center
  - Research recommendations - letters
  - Research policy updates
  - Webinars and policy briefs
  - https://www.ncai.org/policy-research-center/research-data/research-recommendations

- US Census Bureau
  - Disclosure Avoidance System
  - https://www.census.gov/about/policies/privacy/statistical_safeguards/disclosure-avoidance-2020-census.html

- Questions? Email research@ncai.org