Learning about Homelessness in the U.S. Using Linked Administrative and Survey Data

GSS/SSS/ASA Practicum on Blended Data
September 24, 2020

Bruce D. Meyer, Angela Wyse, Alexa Grunwaldt, Carla Medalia, and Derek Wu
Disclaimer

- Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results were approved for release by the Census Bureau, authorization numbers CBDRB-FY20-ERD002-004 and CBDRB-FY20-ERD002-007.
Background

- Part of Comprehensive Income Dataset (CID) project which combines survey and admin data to improve income estimates

- Official poverty statistics and extreme poverty studies are not intended to represent people experiencing homelessness

- People experiencing homelessness not covered or sharply under-represented in most surveys
  - Not generally surveyed in CPS and SIPP; ACS includes only those in shelters

- We use restricted survey and administrative data to:
  - improve income estimates
  - understand survey coverage
  - learn about homeless population
What we hope to learn

Income and Program Receipt
- Employment and earnings in formal labor market
- Safety net program receipt
- Permanence or transience of low material well-being among those who experience homelessness
- Implications for official statistics of the omission of those experiencing homelessness

Other Aspects of this Project
- Population estimates and survey coverage
- Population characteristics
- Transitions between housing states
- Mortality
- Migration and geographic dispersion
Challenges to studying homelessness

- People experiencing homelessness are difficult to survey
  - Reasons include mobility, lack of a permanent residence, tenuous attachment to living quarters, not wanting to be found, pretending to be housed (Glasser, Hirsch, and Chan 2014) or cognitive challenges
  - Raises questions about the representativeness and comprehensiveness of any data source

- There are many different definitions of homelessness
  - We focus on individuals residing in emergency or transitional shelters (“sheltered homeless”) and those whose primary nighttime residence is a public or private place not meant for human habitation (“unsheltered homeless”)
  - Literature is also concerned with precariously housed and “doubled up”; more complicated to do with current data, for future work (Lee, Tyler, and Wright 2010)
Most available studies on homeless income and program receipt are either localized, outdated, self-reported, or some combination of the three (Burt et al. 1999, Metraux et al. 2018).

Our approach takes advantage of large samples that offer a guide to national homeless patterns, including the unsheltered.

- To date, very few studies on homelessness use the Decennial Census and ACS.

We rely on accurate administrative data as well as self-reports.

By linking tax and program data, we get a more detailed picture of situation of those experiencing homelessness including longitudinal information.
Data Sources and Linkage
Census Bureau Data

- **2010 Decennial Census**
  - Enumerated individuals at emergency and transitional shelters, as well as unsheltered individuals in soup kitchens, regularly-scheduled mobile food vans, and targeted non-sheltered outdoor locations (TNSOLs)
  - Enumeration frame developed by internet research and querying local officials, followed by validation and advance visits

- **American Community Survey**
  - Collects micro-level data on individuals in emergency and transitional shelters since 2006
  - Draws on the shelter list from the Decennial, which was expanded starting in 2011
Administrative Income/Resource Data

We link these sources to the following longitudinal administrative data:

- Taxable Income (IRS 1040s, W2s, 1099-Rs)
- Supplemental Nutrition Assistance Program (SNAP) for five states for full time period
  - Illinois, Indiana, New York, New Jersey, and Tennessee
- Medicare and Medicaid enrollment, VA benefits
- Housing assistance (HUD PIC and TRACS)
- Birth and death dates (Numident)
We link the 2010 Census and ACS to administrative tax data using Protected Identification Keys (PIKs) created by the Census Bureau’s Person Identification Validation System (PVS)

- PVS uses name, date of birth, sex, and address information from the Census and ACS to search for a matching record in a reference file derived Social Security Administration (SSA) records
## Linkage (PIK) Rates

### Unweighted Homeless PIK Rates Across Census Bureau Datasets

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS Shelter</td>
<td>0.760</td>
<td>0.736</td>
<td>0.699</td>
<td>0.726</td>
<td>0.827</td>
<td>0.772</td>
<td>0.763</td>
<td>0.774</td>
<td>0.790</td>
<td>0.779</td>
<td>0.750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decennial Shelter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.686</td>
</tr>
<tr>
<td>Decennial Soup Kitchen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.418</td>
</tr>
<tr>
<td>Decennial Food Van</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.424</td>
</tr>
<tr>
<td>Decennial TNSOL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.172</td>
</tr>
<tr>
<td>Houston HMIS¹</td>
<td>0.800</td>
<td>0.949</td>
<td>0.979</td>
<td>0.967</td>
<td>0.955</td>
<td>0.956</td>
<td>0.955</td>
<td>0.961</td>
<td>0.962</td>
<td>0.965</td>
<td>0.965</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.A. HMIS²</td>
<td>1.000</td>
<td>0.895</td>
<td>0.939</td>
<td>0.945</td>
<td>0.870</td>
<td>0.861</td>
<td>0.879</td>
<td>0.906</td>
<td>0.922</td>
<td>0.923</td>
<td>0.925</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sources:** 2006-2016 ACS, 2010 Decennial Census, 2004-2014 Los Angeles CoC HMIS Data, 2004-2014 Houston CoC HMIS Data

**Note:** Table reports the unweighted shares of sheltered and unsheltered homeless individuals who are PIKed in the 2006-2016 ACS and Decennial Census by GQ type. Due to a change in the ACS sampling methodology in 2011 that introduced imputed shelter homeless individuals, we report only the shares of non-imputed shelter homeless individuals who are PIKed from 2011-2016. All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-004.

¹The Houston CoC encompasses shelters in Houston, Harris, Fort Bend, and Montgomery Counties.

²The Los Angeles CoC encompasses shelters in Los Angeles excluding Glendale, Long Beach, and Pasadena.
PIKed vs unPIKed Characteristics

- Characteristics of PIKed and unPIKed homeless individuals in the 2006-2016 ACS
  - Age, gender, race
  - Hispanic origin, born outside the U.S
  - Whether worked in the last year; if report any income receipt

- Many unPIKed individuals are missing sufficient information (name or date of birth) to be PIKed
We adjust for individuals missing PIKs using inverse probability weighting (IPW)

In the Decennial, our model adjusts individual-level weights for the homeless based on:
- Age
- Race
- Gender
- Hispanic origin
- State
- Enumeration type (shelter, soup kitchen, food van, TNSOL)

Covariates are limited in Decennial relative to ACS; may still be some conditional non-randomness in PIKing (especially when PIK rates low)
- We exclude TNSOLs from income and program receipt results due to this concern
Results on Income and Program Receipt
Benefit Receipt

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.

Pre-Tax Income


All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Pre-Tax Income (75th Percentile)


All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
All results were approved for release by the Census Bureau, authorization numbers CBDRB-FY20-ERD002-007.
All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
### Employment & Program Participation Among Adults Ages 18-64

<table>
<thead>
<tr>
<th></th>
<th>Sheltered Homeless</th>
<th>Single Poor Non-Group Quarters</th>
<th>Non-Group Quarters Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worked in Past Year (%)</strong></td>
<td>40.9 (2.09)</td>
<td>45.8 (0.45)</td>
<td>77.6 (0.09)</td>
</tr>
<tr>
<td><strong>Mean Wks Worked in Past 12 Mo. (Cond. On +)</strong></td>
<td>28.4 (1.12)</td>
<td>33.1 (0.23)</td>
<td>44.9 (0.03)</td>
</tr>
<tr>
<td><strong>Mean Hours Worked Per Week (Cond. On +)</strong></td>
<td>32.9 (0.76)</td>
<td>30.2 (0.17)</td>
<td>38.7 (0.03)</td>
</tr>
</tbody>
</table>

### Benefit Receipt Rates and Amounts Among Adults Ages 18-64

<table>
<thead>
<tr>
<th></th>
<th>Sheltered Homeless</th>
<th>Single Poor Non-GQ</th>
<th>Non-GQ Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Retirement or Pension Income Receipt Rate (%)</strong></td>
<td>2.7 (0.78)</td>
<td>2.4 (0.13)</td>
<td>4.5 (0.04)</td>
</tr>
<tr>
<td><strong>Medicaid Receipt Rate (%)</strong></td>
<td>46.7 (2.37)</td>
<td>39.1 (0.49)</td>
<td>10.4 (0.07)</td>
</tr>
<tr>
<td><strong>Food Stamp Receipt Rate (%)</strong></td>
<td>60.4 (1.85)</td>
<td>53.7 (0.46)</td>
<td>13.6 (0.10)</td>
</tr>
<tr>
<td><strong>SSI Receipt Rate (%)</strong></td>
<td>8.8 (0.84)</td>
<td>10.8 (0.30)</td>
<td>2.7 (0.04)</td>
</tr>
<tr>
<td><strong>Public Assistance Receipt Rate (%)</strong></td>
<td>19.2 (1.74)</td>
<td>8.5 (0.25)</td>
<td>1.8 (0.03)</td>
</tr>
<tr>
<td><strong>Received Any Transfer Income (%)</strong></td>
<td>66.3 (1.64)</td>
<td>57.0 (0.47)</td>
<td>15.4 (0.10)</td>
</tr>
<tr>
<td><strong>Received Any Retirement or Employment Income (%)</strong></td>
<td>43.2 (1.93)</td>
<td>47.6 (0.45)</td>
<td>79.7 (0.09)</td>
</tr>
<tr>
<td><strong>Other Income Receipt³ (%)</strong></td>
<td>11.9 (1.17)</td>
<td>12.9 (0.30)</td>
<td>8.6 (0.06)</td>
</tr>
<tr>
<td><strong>Any Income Receipt (%)</strong></td>
<td>72.7 (1.56)</td>
<td>74.1 (0.40)</td>
<td>88.3 (0.08)</td>
</tr>
<tr>
<td><strong>Mean Total Income Amount ($) (Cond. On +)</strong></td>
<td>9,474.0 (470.0)</td>
<td>7,526.0 (55.7)</td>
<td>40,400.0 (129.30)</td>
</tr>
</tbody>
</table>

### Ages 18-64

<table>
<thead>
<tr>
<th></th>
<th>Sheltered Homeless</th>
<th>Single Poor Non-Group Quarters</th>
<th>Non-Group Quarters Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ages 18-64</strong></td>
<td>1,900</td>
<td>17,000</td>
<td>264,000</td>
</tr>
</tbody>
</table>

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Pre-Tax Income by Sub-Group

Median Pretax Income in Administrative Data, 2010 (Ages 18-64)


All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Benefit Receipt by Sub-Group

**SNAP Receipt, 2010 (Ages 18-64)**

- Male Unsheltered: 0.00
- Male Sheltered: 0.02
- Female Unsheltered: 0.01
- Female Sheltered: 0.04
- White Unsheltered: 0.03
- White Sheltered: 0.05
- Black Unsheltered: 0.04
- Black Sheltered: 0.02

**VA Benefit Receipt, 2010 (Ages 18-64)**

- Male Unsheltered: 0.00
- Male Sheltered: 0.02
- Female Unsheltered: 0.01
- Female Sheltered: 0.04
- White Unsheltered: 0.03
- White Sheltered: 0.05
- Black Unsheltered: 0.04
- Black Sheltered: 0.02

**HUD Benefit Receipt, 2010 (Ages 18-64)**

- Male Unsheltered: 0.00
- Male Sheltered: 0.02
- Female Unsheltered: 0.01
- Female Sheltered: 0.04
- White Unsheltered: 0.03
- White Sheltered: 0.05
- Black Unsheltered: 0.04
- Black Sheltered: 0.02


All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Medicare and Medicaid by Sub-Group


All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Overall Impressions of Results

- Lack of employment and reliance on safety net persistent
- Homeless almost all reached by some safety net program
- Administrative data indicate higher rates of SNAP and Medicaid receipt than self-reported in surveys
- Data present new puzzles on differences in income by race, gender, and sheltered/unsheltered status
Implications for Poverty Statistics

- The homeless population is very small relative to the broader population in poverty
  - Including the ACS sheltered homeless in official statistics would increase the poverty rate by between 0.05 and 0.10 percentage points on a base of about 15 percent (15.1 in the 2010 CPS)
  - If we assume all those in the PIT count were poor (sheltered and unsheltered), we would add to poverty between 0.15 and 0.20 percentage points

- Effect on poverty statistics in certain geographic areas or for certain sub-groups (e.g. veterans, people with disabilities) may be more pronounced - a topic we will explore in future work

- Important to look at homeless population separately to understand the deprivation they face


Work Cited


Overall Population Estimates: Sheltered

*Indicates data obtained from publicly available sources.

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-004.
Overall Population Estimates: Unsheltered

*Indicates data obtained from publicly available sources.

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-004.

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-004.
Sheltered Homeless Characteristics

- Share Hispanic and male across data sources
  - About 17% Hispanic
  - About 62% male

- Sheltered homeless by age in HMIS data
  - Share under 18 is 22% in HMIS, 20% in Decennial, 15% in ACS
  - Modal age category in all sources is 31-50 years (about 36% of sheltered homeless)
### Demographics relative to comparison groups

<table>
<thead>
<tr>
<th>All Ages</th>
<th>Sheltered Homeless Mean</th>
<th>Sheltered Homeless SE</th>
<th>Single Poor Non-GQ Mean</th>
<th>Single Poor Non-GQ SE</th>
<th>Non-GQ Mean</th>
<th>Non-GQ SE</th>
<th>NSHAPC (1996)* Mean (Ages 17+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5</td>
<td>5.8 (0.87)</td>
<td></td>
<td>12.3 (0.22)</td>
<td></td>
<td>6.8 (0.05)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>5-17</td>
<td>9.6 (1.32)</td>
<td></td>
<td>25.9 (0.29)</td>
<td></td>
<td>17.9 (0.06)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>18-24</td>
<td>11.1 (1.02)</td>
<td></td>
<td>12.5 (0.25)</td>
<td></td>
<td>9.2 (0.05)</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>25-44</td>
<td>30.2 (1.63)</td>
<td></td>
<td>23.5 (0.28)</td>
<td></td>
<td>26.6 (0.08)</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>45-64</td>
<td>39.9 (2.13)</td>
<td></td>
<td>17.1 (0.23)</td>
<td></td>
<td>26.6 (0.08)</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>&gt;64</td>
<td>3.5 (0.72)</td>
<td></td>
<td>8.7 (0.15)</td>
<td></td>
<td>12.9 (0.05)</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Male (%)</td>
<td>62.1 (2.60)</td>
<td></td>
<td>41.1 (0.32)</td>
<td></td>
<td>48.8 (0.08)</td>
<td></td>
<td>68</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>47.0 (2.54)</td>
<td></td>
<td>56.3 (0.36)</td>
<td></td>
<td>75.5 (0.09)</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Black</td>
<td>40.6 (2.03)</td>
<td></td>
<td>30.7 (0.31)</td>
<td></td>
<td>12.8 (0.07)</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>AI/AN</td>
<td>2.7 (0.52)</td>
<td></td>
<td>2.2 (0.09)</td>
<td></td>
<td>1.2 (0.02)</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Asian</td>
<td>1.8 (0.49)</td>
<td></td>
<td>2.3 (0.10)</td>
<td></td>
<td>5.1 (0.04)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>7.9 (1.13)</td>
<td></td>
<td>8.5 (0.22)</td>
<td></td>
<td>5.4 (0.05)</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Hispanic (%)</td>
<td>17.0 (1.42)</td>
<td></td>
<td>25.0 (0.30)</td>
<td></td>
<td>17.0 (0.07)</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Total Sample Size</td>
<td>2,300</td>
<td></td>
<td>32,500</td>
<td></td>
<td>433,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pop. Estimate/Weighted Count</td>
<td>165,400</td>
<td></td>
<td>2,684,000</td>
<td></td>
<td>30,150,000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sources:** 2010 American Community Survey, Burt et al. 1999

**Note:** The ACS characteristics for GQ individuals are computed among non-imputed individuals using survey weights that are scaled up by a constant such that the new weighted count of non-imputed observations is equal to the old weighted sum of imputed and non-imputed records. * Indicates data from publicly available sources.

All results were approved for release by the Census Bureau, authorization number CBDRB-FY20-ERD002-007.
Recent Trends in California and NY


Note: All data obtained from publicly available sources. San Francisco Bay includes the following CoCs: San Francisco, Contra Costa County, Marin County, Alameda County, San Mateo County, and Santa Clara County. LA, OC, and SD includes: Los Angeles, Pasadena, Glendale, Long Beach, Orange County, and San Diego.
Recent Rises in California and NY

California PIT Estimates

New York City PIT Estimates


Note: All data obtained from publicly available sources. San Francisco Bay includes the following CoCs: San Francisco, Contra Costa County, Marin County, Alameda County, San Mateo County, and Santa Clara County. LA, OC, and SD includes: Los Angeles, Pasadena, Glendale, Long Beach, Orange County, and San Diego.