



M William Sermons, PhD

Vice President, Research, Volunteers of America
Adjunct Professor, Carnegie Mellon University
Advisor, Eviset

Carnegie Mellon University
HeinzCollege



Selected Federal Data Sources by Type

Census and Population Data	Housing Data	Economic Data	Administrative Data	Health Data
<ul style="list-style-type: none">• Decennial Census• American Community Survey• Current Population Survey	<ul style="list-style-type: none">• American Housing Survey• Housing Inventory Count (HIC)	<ul style="list-style-type: none">• Survey of Income and Program Participation• Consumer Expenditure Survey	<ul style="list-style-type: none">• National Provider Identifier (NPI) Registry• Federal Audit Clearinghouse	<ul style="list-style-type: none">• National Health Interview Survey• National Health and Nutrition Examination Survey

Organization-level federal data sources.

Leveraging Organization-Level Federal Data to Understand Organizations/Communities

Using organization-level data can facilitate:

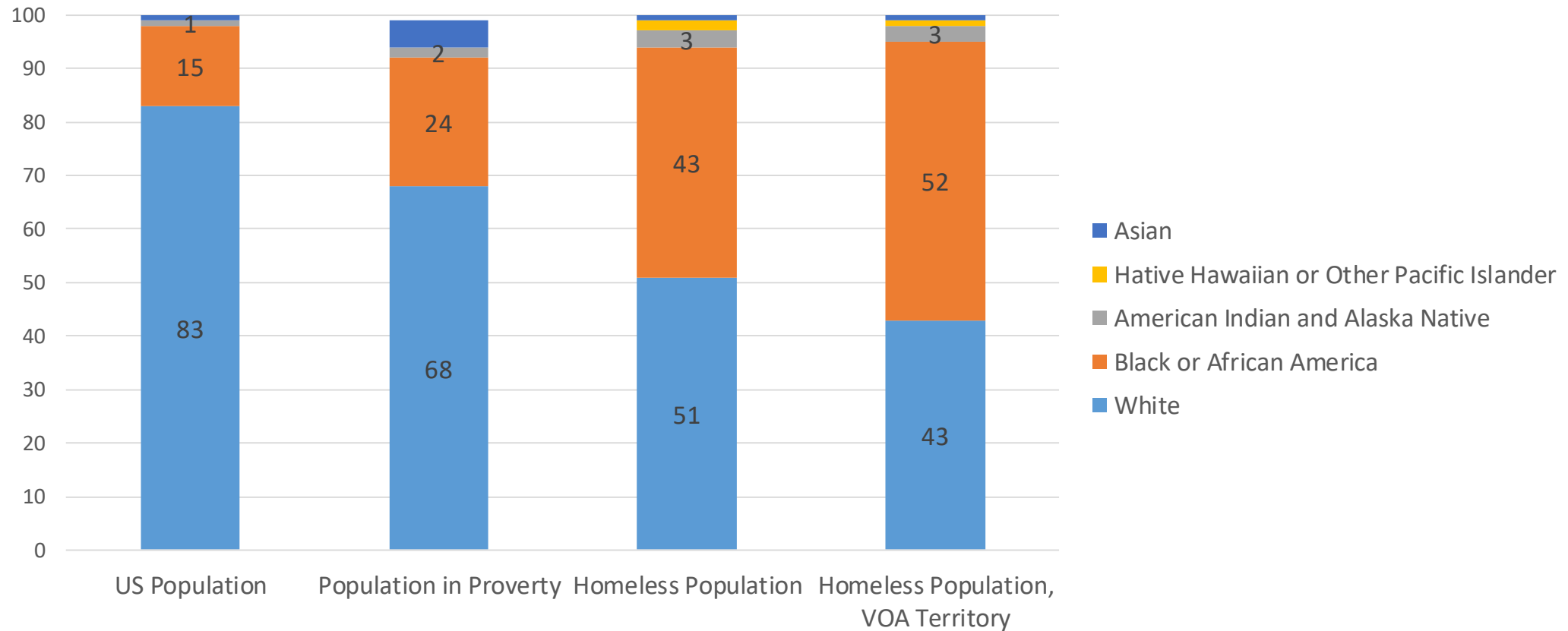
- Understanding of organizational actors central to policy/practice analysis
- Increased data consistency across entities
- Significant reduction in data collection costs and effort
- Comparisons across markets and communities

Considerations When Using Organization-Level Federal Data

- Quality control challenges with regulatory compliance data
- Size heterogeneity present presentation or analysis challenges
- Data appropriateness assessment
- Complex identifier management
- Industry-Specific knowledge may be needed

Examples

Blending Population & Homeless Assistance Application Data to Understand Affected Populations



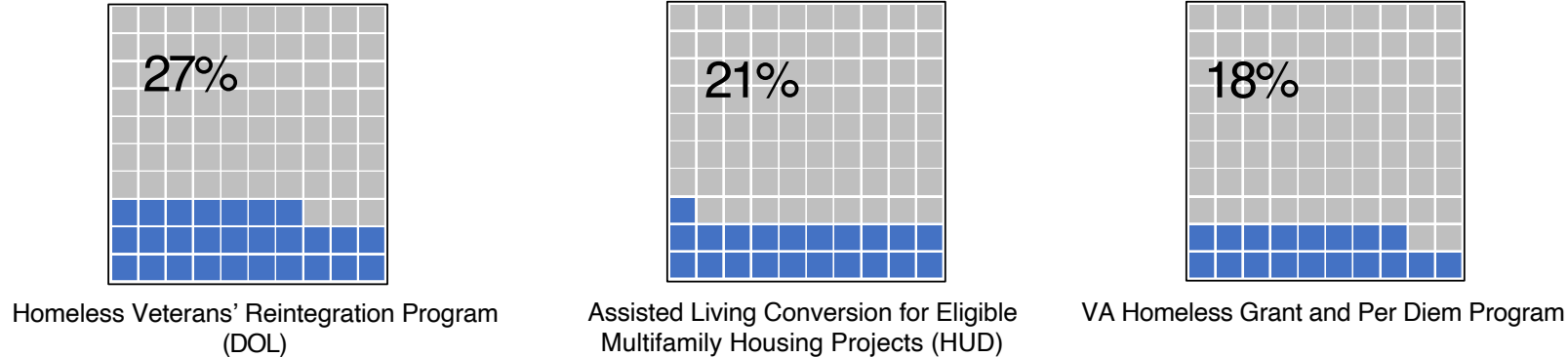
Sources: Census Current Population Survey Data and Continuum of Care Data. This CoC data is based on point-in-time information provided to HUD by Continuums of Care (CoCs) as part of their CoC Program application process, per the Notice of Funding Availability (NOFA) for the Fiscal Year 2022 Continuum of Care Program Competition.

Using Federal Expenditure Data from Audit Clearinghouse to Estimate Market Share

VOA Expenditures vs. Total Expenditures

VOA
125

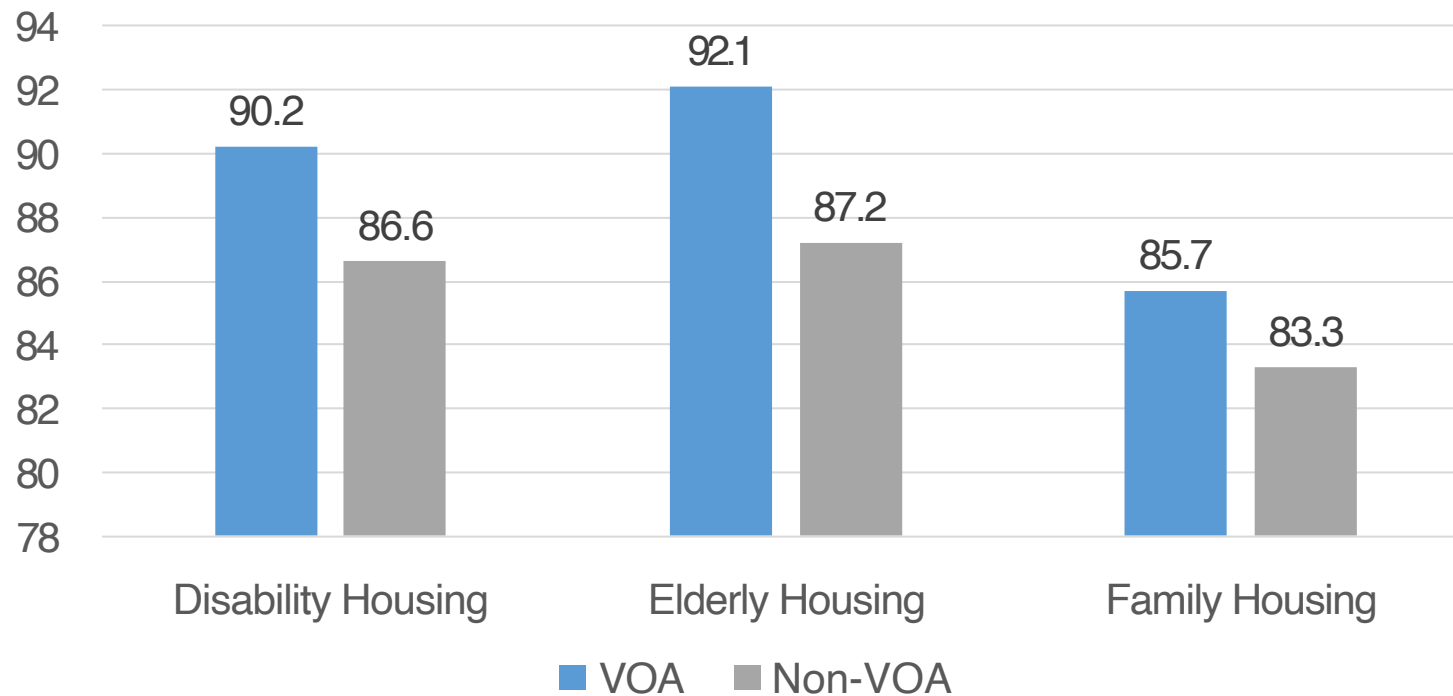
VOA Fraction of Total Program Expenditures(%)



Sources: Federal Audit Clearinghouse of the Office of Management and Budget. Single audits of entities that receive and expend federal awards.

Using Real Estate Assessment Data for Market/Peer Quality Comparisons

Comparing Publicly Supported VOA Housing to Publicly Supported Non-VOA Housing



HUD's Real Estate Assessment Center conducts physical property inspections of properties that are owned, insured or subsidized by HUD, including public housing and multifamily assisted housing. About 20,000 such inspections are conducted each year to ensure that assisted families have housing that is decent, safe, sanitary and in good repair.

Sources: The data in the NHPD come from the US Department of Housing and Urban Development (HUD) and the US Department of Agriculture (USDA), and include ten federally subsidized programs.

Medicaid Shiny Dashboard

The NPI (National Provider Identifier) registry is a database maintained by the Centers for Medicare & Medicaid Services (CMS) that contains information about healthcare providers in the United States. Covered healthcare providers, health plans, and healthcare clearinghouses must use NPIs in administrative and financial transactions adopted under HIPAA. The registry allows you to search for information about healthcare providers, including organization names, doctor names, taxonomy codes, specializations, addresses, and more.

Create R code to identify locations of multiple nonprofit organizations (Catholic Charities, Lutheran Services, Salvation Army, Volunteers of America, YMCA, and YWCA) that can bill Medicaid, and for which service types (taxonomy).

Dashboard Steps

1. Query the NPI Registry API (<https://npiregistry.cms.hhs.gov/api/>) to get NPI data for these organizations
2. Handle the API's 200-record pagination limit properly
3. Process the nested structure of API responses to extract:
 1. Organization name
 2. Location details (address, city, state, zip)
 3. Taxonomy codes and descriptions
 4. Medicaid billing capability

Dashboard Steps

4. Add an organization identifier column to track which organization each record belongs to
5. Geocode the addresses using the Census API
Organization name
 - Location details (address, city, state, zip)
 - Taxonomy codes and descriptions
 - Medicaid billing capability
6. Create a Shiny dashboard that allows users to: -
 - Filter by organization, state, and taxonomy
 - View locations on an interactive map with clustering

NPI Data Acquisition Using R

```
33 # Define base URL for NPI Registry
34 base_url <- "https://npiregistry.cms.hhs.gov/api/?version=2.1"
35
36 # Function to get NPI data for an organization with a skip value (pagination)
37 get_npi_data <- function(org_name, skip_val = 0) {
38   # Format the organization name for the API query
39   # Adding wildcard to capture variations in name
40   formatted_org_name <- paste0(org_name, "*")
41
42   # Set parameters for the API request
43   params <- list(
44     organization_name = formatted_org_name,
45     enumeration_type = "NPI-2", # Organization NPI type
46     limit = 200,                # Maximum records per request
47     skip = skip_val             # Pagination offset
48   )
49
50   cat(paste0(" Querying API with: ", formatted_org_name, " (skip=", skip_val, ")\n"))
51
52   # Make the API request
53   response <- GET(base_url, query = params)
```

Conclusion

- The wide availability of household and individual survey data often overshadows valuable entity-level government administrative data as an important research resource.
- Organizations, researchers, and analysts can leverage these administrative data sources to gain unique insights into organizational activities, market positions, and comparative performance.
- When using administrative data sources, it's essential to address their limitations, including data quality variability, organizational size differences, and the management of various identification systems (EINs, DUNS, etc.).