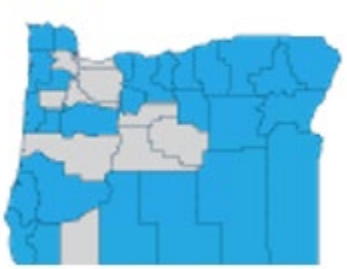
**Evaluating a Sampling Method for Completing the Point-in-Time Count:**

**A Case Study of the 2022 Rural Oregon Continuum of Care**

## Introduction

In the early 2000s, the United States Department of Housing and Urban Development (HUD) initiated the Point-in-Time (PIT) Count, a biennial count of sheltered and unsheltered homeless persons on a single night, the last Wednesday in January. The PIT Count is conducted by HUD-defined Continuums of Care (CoC) which are regional planning bodies that coordinate housing and services for homeless families and individuals.

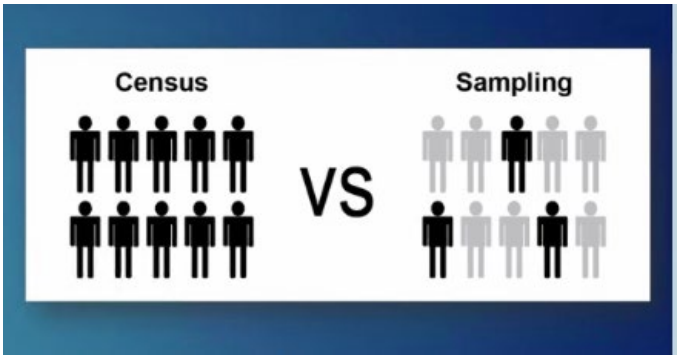
In Oregon, the Rural Oregon Continuum of Care (ROCC) is responsible for the PIT Count in 26 counties, served by six regional community action agencies.



In the fall of 2021, Community Services Consortium (CSC), the community action agency for Benton, Linn, and Lincoln counties, asked the Oregon State University Policy Analysis Lab (OPAL) to assist the CSC and ROCC with planning and conducting the PIT Count using a sampling methodology permitted and promoted by HUD. Jurisdictions are permitted to count unsheltered homeless people for the last 10 days of January. The data for sheltered homeless were pulled on January 26, 2022.

## Methodology

HUD provides guidelines for conducting the PIT Count through both a census and sampling methodology. In both approaches, sheltered homeless numbers are obtained from the Homeless Management Information System (HMIS) and included in the PIT Count. Counting unsheltered homeless populations is done by volunteer enumerators. Volunteer enumerators are usually employees of agencies that serve homeless people, street outreach professionals, and community members. The *census method* for counting unsheltered homeless involves covering




the region with volunteers seeking to count all homeless individuals throughout the entire geography of the CoC's service region. The *sampling method* involves dividing the CoC's service

region into U.S. Census tracts (or some other sub-county area) from which a portion of them are selected and then the number of unsheltered is obtained from only the sampled areas and extrapolated to estimate the total unsheltered homeless population in the service region.

The sampling approach offers the benefit of permitting CoCs to focus resources on counting in a smaller number of areas within their service region. Deciding which approach to use involves tradeoffs, but limiting factors for most CoCs are people and time. As such, sampling is an attractive option as it can generate reliable estimates of the homeless population and requires fewer resources for CoCs that are already stretched thin.

In both methodologies (census or sampling), counts are conducted using either surveys of individuals, which is a set of approximately 30 questions, or through observations made by volunteer enumerators, assessing various characteristics of persons appearing to be in the unsheltered homeless category. Surveys provide more detailed information about an individual's characteristics but require significantly more time and rely on interviewees providing accurate information about themselves. Observational counts allow volunteers and staff to count more people more efficiently but rely on the judgment of those doing the count to determine whether someone is homeless. Moreover, observational studies do not generate trustworthy demographic information about individuals. We do not here test the relative merits of these two data collection strategies, but simply note again that a COC's capacity likely will shape this decision.

The focus of this report is on the practice and value of the sampling method. This method can be divided into two phases. Phase one involves designating some areas as high probability areas with known large populations of homeless individuals that warrant a full census-like counting effort and designating other areas as lower probability areas from which to sample. To make this determination, a threshold value is first generated. This value is informed by using data from previous PIT Counts and local knowledge. Any tract or area where previous counts were high, or where local officials know homeless counts are higher than this threshold value, are qualified as high probability areas, and receive most of the counting efforts by enumerators. Any tract where previous counts were lower than the threshold value or local officials know there are not many homeless individuals are designated as low probability areas and are part of the pool of areas from which a sample of tracts is taken. Thus fewer volunteer resources are allocated to these relatively few tracts with anticipated low numbers of homeless persons.

Phase two involves randomly selecting from the designated low probability sampling areas, then assigning enumerators to count only within those selected areas. After this is completed, the number generated from the count in the selected areas is multiplied by a weighting factor. Thus, the low probability sampling areas that were not chosen are assigned an estimated homeless count based on the weighting factor and on the count in the areas that were chosen. The weighting factor is determined by dividing the number of low density tracts by the total number of selected low probability tracts where a count was conducted. This weighting factor is multiplied by the total number of unsheltered homeless counted in the selected tracts. The total number of unsheltered homeless persons using this method is generated by combining the

sampling estimate with the census counts from the high probability areas. We illustrate this method with real data from the 2022 PIT count completed for the ROCC.

OPAL collaborated with the ROCC to designate low probability census tracts using data from the 2020 PIT count. We determined that U.S. Census tracts where 10 or less people were counted in 2020 would be designated as low probability areas. This resulted in identifying 111 low probability tracts, of which 22 were randomly selected to be sampled. This generated a weighting factor of 5.045 (111/22) meaning that the total number of individuals counted within the total 22 tracts selected would be multiplied by 5.045 and the resulting product would be the estimated number of unsheltered homeless individuals within all the low probability tracts in the ROCC region.

An important contextual note is in order for interpreting the 2022 data. The week prior to the scheduled count, HUD notified the CoCs that the January 2022 count could be postponed to February due to the Omicron surge. The ROCC chose to go ahead with the plans for January, and many volunteers and staff had to cancel their participation due to illness. Additionally, although OPAL was collaborating with the ROCC to set up a sampling methodology, only half of the community action agencies opted-in to the pilot. The disorganized approach to the 2022 count methodology provided data that is of questionable validity in terms of accurately reporting the full number of unsheltered homeless; however, fortuitously the data can be used to demonstrate the possible usefulness of sampling. We therefore illustrate how the numbers generated from the count as it was done differ from what they could have been if sampling were relied on. Indeed, the results show here that sampling provides a useful and HUD-authorized strategy that CoCs could use under difficult circumstances such as those experienced in 2022 when time and resources were especially constrained such that counts of unsheltered homeless would be underestimated.

**Results**

A total of 3,305 sheltered and unsheltered homeless individuals were counted throughout the ROCC through the 2022 PIT count. Out of this total, 2,894 were unsheltered. While 90% of these were counted in the high probability tracts, where census-counting efforts were likely concentrated, 296 of these unsheltered individuals (about 10%) were located in tracts designated as low probability (regardless of whether those tracts were chosen in the sampling effort).

In the sampled low probability tracts, 126 unsheltered homeless individuals were counted. Following the method described above, beginning with the weighting factor of 5.045 (111/22), and multiplying this by 126, we would estimate that in the low probability tracts of the ROCC there were 126 x 5.045 = 635 unsheltered homeless persons. This is a much higher number than the 296 counted in all the low-probability tracts that were counted. In other words, this is an undercount of 339 persons in the low probability tracts. This represents an undercount of around 10% of all unsheltered homeless in the ROCC [339/(339+2894)]. Not every sampled tract was counted by each CoC, thus underestimating the number of unsheltered homeless in selected tracts, and thus increasing the size of the undercount. The size of the undercount cannot be precisely known, but this initial observation suggests that sampling could easily increase the ROCC's estimate of unsheltered homeless by at least 10%.

To further demonstrate the county-specific benefits of sampling we calculated the expected number of persons in each census tract by simply dividing the 126 persons by the 22 tracts for an estimate of 5.73 persons per low probability tract. Each county has a unique number of low-probability U.S. Census tracts. Multiplying the number of low probability tracts by the estimated average of unsheltered homeless per low probability tract, we estimated how many unsheltered homeless persons per county would have been reported versus the actual numbers collected by volunteers for the 2022 Pit Count. Table 1 compares each county's estimated count for low probability areas with the actual count within low probability census tracts in those counties. The numbers show that the actual counts obtained by volunteers likely undercounted in almost every county, with the exception of Linn and Josephine counties.

Table 1. Comparing Estimated Low Probability Census Tract Estimates of Unsheltered Homeless with Actual Counts, by County

| County | Estimated County Totals (with sampling) | County Totals Obtained by Volunteers (no sampling) |
|---|---|---|
| Linn | 80.2 | 88 |
| Yamhill | 74.5 | 40 |
| Josephine | 40.1 | 40 |
| Douglas | 51.5 | 36 |
| Columbia | 34.4 | 30 |
| Klamath | 40.1 | 14 |
| Lincoln | 51.5 | 11 |
| Clatsop | 40.1 | 7 |
| Malheur | 34.4 | 7 |
| Hood River | 11.5 | 7 |
| Union | 22.9 | 6 |
| Grant | 11.5 | 6 |
| Benton | 57.3 | 4 |
| Tillamook | 28.6 | 0 |
| Baker | 22.9 | 0 |
| Coos | 11.5 | 0 |
| Grant | 11.5 | 0 |
| Wallowa | 11.5 | 0 |
| Harney | 5.7 | 0 |
| Umatilla | 5.7 | 0 |
| Lake | 5.7 | 0 |
| | | |
| Total | 647.2 | 296 |

Table 1 shows that our estimate was higher in every county except for Linn County and Josephine County. Additionally, as in our above-demonstrated analysis, our estimate produced a total that is over 300 persons undercounted. This table also shows that some counties, as far ranging as Benton, Malheur or Clatsop counties, may be substantially underestimating the number of unsheltered homeless (25 to 50 persons) in the low-probability areas of their county. We are cautious in asserting that any one county is most at risk of undercounting, especially when applying regional averages to the counties that comprise them, but the apparent variation suggests that counties and the community action agencies that serve there could benefit from re-examining how their low probability areas may be unexpectedly populated by small but not insignificant numbers of unsheltered homeless. Of particular import is the likelihood that nearly all counties contribute to an undercount of unsheltered homeless in the ROCC.

**Conclusion**

Our investigation of the 2022 PIT count data for the ROCC provides an example of how sampling could have impacted the 2022 PIT count and how useful it could be for future PIT counts. Notably, counties that did not have the capacity to conduct counts in low probability areas could have benefitted from sampling to potentially produce a more accurate count of the homeless population for the whole ROCC. Additionally, counties that were able to send teams out to many of their low probability tracts could have used sampling to increase the efficiency and accuracy of their counts and allocate other teams to more high probability areas. Our analysis shows that not using a sampling method to estimate the unsheltered homeless population size for the ROCC likely underestimated the county by at least 10%, and that this undercount was widespread throughout the ROCC service area.

It is important to understand that because of the cross-sectional nature of the PIT Count, and HUD's definitions of homelessness, the PIT Count always undercounts the number of homeless persons. Sampling cannot address the strict federal definitions of sheltered and unsheltered people that contribute to ignoring populations of homeless people such as those taking shelter with friends and family. But it appears that sampling may help counteract undercounting by at least addressing the resource and volunteer limitations of CoCs. It is important to note that PIT Count and sampling is not the only means by which community leaders can understand the geography and scope of the problem of homelessness. Other on-going methods throughout the year, and more effective use of administrative data collected as homeless individuals access services, can validate survey results such as the PIT count. Because the annual PIT count will likely continue as a required effort for CoCs, it is valuable to consider how familiar counting efforts may fall short in producing an accurate estimate, while also over-stretching CoCs trying to complete the PIT count. Sampling offers an opportunity to obtain a more valid estimate that takes seriously staffing and resource limitations. There are tradeoffs associated with choosing any methodology but sampling provides reliable and trustworthy counts using fewer resources.

This report prepared by OSU Policy Analysis Laboratory (OPAL) researchers.

Mark Edwards is Professor of Sociology in the School of Public Policy, and Director of OPAL.
Catalina Harmon is a 2nd year Master of Public Policy student.
Sabrina Beck is a recent graduate of OSU's Master of Public Policy program.

October 16, 2022