Calculating County-Level Housing Choice Voucher Gaps: A Methodology

Shane Dabney
Florida State University

Abstract

This article presents a methodology for calculating county-level housing choice voucher gaps, leveraging IPUMS-obtained 5-year American Community Survey (ACS) data. Using Florida as a case study, the approach builds on existing methodologies to offer a fine-grained calculation that can be readily adapted to other states, providing a more precise quantification of voucher needs. Such precision is essential for informing local, state, and federal housing policies and enhancing the accuracy of research relying on precise counts of voucher eligibility at the local level. The methodology utilizes custom Python algorithms to prepare the ACS data by imputing missing county values and accounting for relevant factors, such as multifamily households and incarcerated persons, in the final eligibility determination (Dabney, 2024). Applying this methodology to Florida reveals significantly wider voucher gaps than common national estimates suggest, indicating a more substantial discrepancy between the number of families in need and those currently receiving vouchers than previously recognized.

Introduction

The number of Housing Choice Vouchers (HCVs) distributed nationally is considerably less than the number of voucher-eligible households. This disparity between eligible families and available aid is known as the *voucher gap*. Accurately measuring the scale of this gap is crucial for policymakers to evaluate the voucher program's coverage and effectiveness.

Although the national and state-level voucher gaps are often discussed, more fine-grained gap calculations, such as at the county level, are rare. However, such estimates have many potential uses. For policymakers, more granular voucher gap data can enable precise comparative analysis, offering insights into regional disparities within a state. For program practitioners, accurate estimates of these gaps can aid program calibration or demonstrate the need for higher levels of funding.

No standardized methodology currently exists for calculating voucher gaps, especially at localized levels. A widely cited national estimate is that only about one in four eligible households receive vouchers. However, this baseline figure glosses over geographic variability and methodological limitations. Developing a more rigorous approach is critical as local public housing authorities make decisions about voucher allotments and eligibility requirements.

This article introduces a refined methodology for calculating voucher gaps at the county level. The approach utilizes custom Python algorithms to prepare IPUMS-obtained American Community Survey (ACS) data, imputing missing county values and accounting for various household compositions, including multifamily households and incarcerated persons (Dabney, 2024; IPUMS, n.d.). Although the methodology is applied specifically to Florida, the framework is generalizable and should be applicable to other states with only minor required variations. The approach promises a more precise gap estimate, helping policymakers to pinpoint priority areas for additional vouchers or eligibility adjustments.

The results of this calculation in Florida starkly illustrate the inadequacy of current voucher allocations relative to need. Across the state, no county reaches the often-cited 25 percent national allocation rate. In fact, the highest county allocation rate in Florida barely approaches 10 percent, with several counties languishing at a mere 1 percent. This significant discrepancy underscores the urgent need for policymakers to address the vast disparity between the availability of vouchers and the actual demand for housing assistance.

Data Sources and Methodology Overview[1]

1. IPUMS-obtained 5-year ACS data were used to determine county, income, family size, and various household characteristics useful for additional research (Ruggles et al., 2023).

2. Several ACS records lacked county values. Nonetheless, because all records contained Public Use Microdata Area (PUMA) values, these were used to fill in the missing county values. This was achieved using crosswalk data obtained from Geocorr 2018: Geographic Correspondence Engine (MCDC, 2018).

3. A custom algorithm then checked and cleaned income data, accounting for distinct family incomes in multifamily households.

4. Multifamily households were then split into separate households. This step allows eligibility determinations at the family level rather than at the household level for multifamily households. This step necessitates a recalculation of the household weight (HHWT) variable.

5. An aggregation algorithm was then used to flatten households into a single row and aggregate various family characteristics. This step is optional because these family characteristics are not strictly needed to calculate voucher gaps.

[1] It should be noted that although some important adjustments have been made, this methodology was inspired by and closely follows the voucher gap calculation methodology detailed in The Housing Initiative at Penn's report (Reina, Aiken, and Epstein, 2021).

6. Eligibility was then calculated by comparing each household's income and family size with U.S. Department of Housing and Urban Development (HUD) income limits for the relevant areas (HUD, n.d.a.).

7. Because ACS data include incarcerated persons, these individuals were removed from the final eligibility count. This step involved importing prisoner data from the Florida Department of Corrections and the Federal Bureau of Prisons.

8. Finally, voucher gaps and allocation rates were determined by comparing each county's final eligibility count with the number of vouchers allocated for those counties. These data are from HUD's Picture of Subsidized Households (HUD, n.d.b.).

Methodology

Imputation of Missing County Data

Approximately 20 percent of the 2021 ACS Florida records were missing county data. Dropping such a substantial percentage of rows was not a viable option. Fortunately, no records were missing PUMA values, which could be cross-referenced to county values using Geocorr 2018: Geographic Correspondence Engine (MCDC, 2018). After creating a data frame to match PUMA codes to counties, a merge function was used to impute the missing county data in the ACS dataset.

Unfortunately, some PUMAs cover more than one county. Even so, given the absence of a more granular geographic variable in the dataset that did not present similar overlap issues, PUMA values were used to perform the crosswalk. The relevant counties were combined into larger geographic county groups to address the overlap, which was necessary for several Florida counties. This approach retained county-level granularity to the greatest extent possible, although it introduced some ambiguity because specific households within overlapping PUMAs could not be pinpointed to a single county.

Alternatively, weights could be applied to allocate households within overlapping PUMAs to individual counties without combining them. This option remains viable for researchers seeking to maintain these county-level distinctions.

Family Size and Income Variables Cleaning

The IPUMS ACS data identifies households by unique serial numbers assigned to each family, the *CBSERIAL* variable. The variable *FTOTINC* represents family income, and *HHINCOME* denotes household income. Discrepancies between these variables can arise in multifamily households—identified when the number of families (NUMFAMS) is more than one—when distinct families have separate incomes.

An accurate voucher gap calculation requires correcting any anomalies in the data with respect to these variables. For instance, Family size (FAMSIZE) should reflect realistic family sizes that are neither negative nor implausibly high. Similarly, income variables should be checked thoroughly. Because the dataset was too large to check manually, it needed to be reviewed and cleaned algorithmically.

Each household's HHINCOME value was compared to the other available income variables for the household members: income from investments (designated with the IPUMS variable INCINVST), retirement (INCRETIR), welfare (INCWELFR), Social Security (INCSS), and Supplemental Security Income (INCSUPP). The total income of each individual within a household is also reported (INCWAGE). All these variables were used to confirm the correct HHINCOME value for each household, which was done with a custom algorithm that summed all relevant individual income values for a given person and cross-checked against the total income recorded in the INCWAGE variable, flagging any anomalies for review. Then, the individual total incomes for each person assigned to the same CBSERIAL were summed and checked against the HHINCOME value. Again, any anomalies were flagged for manual review.

Multifamily Household Split and Income Cleaning[2]

HCVs can be allocated to any household meeting the eligibility criteria. However, it is not uncommon for a household to be comprised of several families living together in the same housing unit. Accounting for these *multifamily* households (NFAMS)—when there is more than one family in a household—is essential for getting an accurate HCV eligibility count because although their combined income might surpass eligibility thresholds, the individual families, if considered separately, could each qualify for a voucher. Similarly, a household eligible as a single entity might actually consist of multiple families, each of which could qualify on their own if assessed independently. Therefore, to avoid undercounting, multifamily households should be split into distinct families and assessed separately.

The ways to split multifamily households into distinct family-level households are numerous. This methodology assigned a new unique serial number to each family. This was done by using the original CBSERIAL value assigned to the household in conjunction with the family unit (FAMUNIT) variable. This FAMUNIT variable identifies the family unit to which a given person in a household belongs. Within single-family households, all individuals have FAMUNIT = 1. For multifamily households with three families, for example, all members of each family are assigned a 1, 2, or 3 in this column. Combining CBSERIAL with the given FAMUNIT value produces a unique family-level (rather than household-level) serial number named FAMILYNUMBER by the algorithm. A final income check was then performed to ensure that the total household income value was not inadvertently used in the eligibility calculation.

This methodological step ensures that independently eligible families in multifamily households are not overlooked. As such, it estimates the total number of voucher-eligible families irrespective of housing composition. It should be noted that, in some cases, multifamily households might choose to continue living together in the same housing unit despite being independently eligible. In these cases, such households would be eligible for only one voucher. This methodology does not account for this possibility, which should be kept in mind when interpreting the voucher gap estimates.

[2] The inspiration for this methodological step comes from The Housing Initiative at Penn's report (Reina, Aiken, and Epstein, 2021).

Household Weight Adjustment

The HHWT variable adjusts for sampling disparities in the ACS data and ensures that each household's data reflects its proportionate presence in the larger population. Correctly applying these weights is essential for producing accurate figures. Per IPUMS' (n.d.) description, "HHWT must be used to obtain nationally representative statistics for household-level analyses of any sample other than [unweighted IPUMS samples]."

It should be noted that algorithmically splitting multifamily households into disparate unit restructures households in a way that is not accounted for in the ACS-assigned weights. If the original household weights are applied to the resulting family-level restructured households, the weights inadvertently will be applied multiple times. To see why, consider a 3-family multifamily household (NFAMS = 3) with a household weight of 15. This household represents 15 other households in the population. However, if the household were to be split and the original household weight preserved, the 3 resulting family-level households would each represent 15 other households, misleadingly suggesting that 45 similar households exist in the population.

Although this increase is appropriate with respect to eligibility determination, it artificially inflates the number of households in the population. For this reason, the household weight of split multifamily households was divided. Specifically, for each split multifamily household, the original HHWT was divided by the relevant NFAMS variable number, effectively dividing the weight equally among the resulting family-level households. In the previous example, each resulting family-level household would have been assigned a new household weight of 5.

$$ADJUSTED\ HHWT = \frac{ORIGINAL\ HHWT}{NFAMS}$$

Optional Aggregation and Flattening

A custom algorithm then aggregated various family characteristic variables and ultimately condensed the data into one row per family. This step was done with the aim of simplifying the data, enhancing its readability, and streamlining subsequent analyses and calculations.

This aggregation grouped data by each unique family unit, summarizing information within these groups. Depending on the data type, the algorithm either took the first value from each group (for consistent family characteristics) or summed values (for relevant quantifiable data). This procedure resulted in a streamlined dataset with each row representing an individual family unit, retaining crucial information in a more manageable form.

Although not strictly necessary for voucher gap calculations, which require only family income and family size data, maintaining detailed fine-grained household characteristics alongside voucher eligibility can be useful for researchers. For example, it allows one to compare voucher-eligible-family characteristics like race with the ineligible population. Likewise, it allows one to compare the demographic characteristics of eligible families with those of actual voucher recipients.

Eligibility Determination

Determining eligibility was then a fairly straightforward process. Each household's location, family income, and family size were matched to HUD's published income limits for the relevant geographic area. These income limits are accessible through HUD's website or its application programming interface (API), with an access token required for the latter (HUD, n.d.a). Although income thresholds for eligibility can technically vary by public housing authority, the general income limit is 50 percent of the Area Median Income (AMI) and, as such, was the one used. Eligibility counts for each county were then compiled into a data frame.[3,4]

Removing Incarcerated Individuals via Stratified Selection

A final consideration involved the housing status of incarcerated individuals in Florida. Incarcerated individuals typically have incomes well below the single-person household limit. Therefore, they will be counted as HCV-eligible by default. However, incarcerated individuals are already housed and not eligible for housing subsidies. Therefore, their eligibility status was adjusted in the dataset.

The ACS data include incarcerated individuals under the group quarters (GQTYPE) variable. Unfortunately, a more specific group quarters type designation was absent from the data than either institutional group quarters (GQTYPE = 1), which includes more than just prisoners, or noninstitutional group quarters (GQTYPE = 5). Without this direct identifier in the ACS data, inmate counts had to be obtained from the Florida Department of Corrections and the Federal Bureau of Prisons, which detailed prisoner counts by facility and county and some basic demographic information, such as race. Similar data may be elsewhere.

The adjustment would be straightforward for a simple eligibility count, only requiring reducing the eligibility count by the number of prisoners in each area; however, attempting to preserve the previously aggregated household characteristics to the greatest possible extent complicated this process. It required rows corresponding to inmates to be identified and selected among the institutional group quarters records.

A county-specific stratified selection technique was employed to approximate this selection. The algorithm targeted rows designated as households in institutional group quarters (GQTYPE = 1) within each county, selecting them to reflect the racial proportions in the incarceration statistics using the RACE variable. The algorithm kept a running total of the adjusted household weight (REALHHWT) during the selection process,[5] which was conducted iteratively until closely aligning with the actual prisoner count for each county. Subsequently, the algorithm updated the relevant eligibility status in the dataset, switching from eligible (1) to ineligible (0) across the income-threshold variables for the relevant rows.

[3] Note that, in some cases, counties assigned to the same county group had different income limits. To address this disparity, the algorithm considered eligibility at both the lowest and highest income limits within the group.

[4] Also, note that some voucher gap calculation methodologies, such as Reina, Aiken, and Epstein (2021), drop homeowners from the final count, which is likely because the HCV program provides rental subsidies, and dropping homeowners aligns with this policy. However, because housing tenure does not affect eligibility according to the Code of Federal Regulations (24 CFR § 982.201), homeowners were retained in this methodology.

[5] The REALHHWT variable name was assigned to the HHWT after the multifamily split and adjustment.

Overall, this approach offered a practical solution given the constraints of the data, but it is important to acknowledge its limitations. Although it preserved the correct racial proportion of HCV-eligible records, it risked slightly skewing other household characteristics. Although preserving the racial proportions was the primary concern here, the technique should be adjusted to reflect the aims of other researchers.[6]

Results

Finalized eligibility counts were compared with the voucher allocation data from HUD's Picture of Subsidized Households and adjusted to reflect the combined county areas (HUD, n.d.c.). The voucher gap in each area is the difference between the calculated eligibility count and the actual subsidy count from HUD. The allocation rate is the percentage of eligible households for which an HCV is actually available (exhibit 1).

Exhibit 1

Florida's County Level Voucher Gaps (1 of 2)

Florida County	Voucher Gap	Allocation Rate (%)
Alachua	54,527	5.3
Brevard	62,518	4.7
Broward	248,899	4.6
Charlotte	18,851	1.7
Citrus	16,520	1.7
Clay	17,674	1.1
Collier	43,395	1.1
County Group 1	19,891	2.4
Duval	123,297	5.7
Escambia	46,527	4.7
Flagler	11,862	2.2
Hernando	25,118	1.7
County Group 2	32,107	0.5
Hillsborough	165,885	6.4
Indian River	15,882	2.5
Jackson	17,845	0.8
Lee	78,230	3.6
Leon	63,964	4.3
Manatee	41,569	4.4
Marion	35,481	4.0
Martin	17,322	0.5
Miami-Dade	406,567	6.8
Monroe	16,156	2.9
County Group 3	9,859	2.2
Okaloosa	24,214	4.2
Orange	173,820	3.1

[6] Alternative approaches focusing only on households or individuals renting a home (OWNERSHP = 2) would inherently exclude incarcerated individuals, obviating the need for this step. However, such an approach risks significantly undercounting eligible households because it does not align with federal eligibility requirements, which do not consider housing tenure.

Exhibit 1

Florida's County Level Voucher Gaps (2 of 2)

Florida County	Voucher Gap	Allocation Rate (%)
Osceola	35,020	0.8
Palm Beach	186,175	4.5
Pasco	66,030	2.7
Pinellas	114,462	7.4
Polk	70,781	3.0
County Group 4	12,999	3.9
Santa Rosa	13,156	2.6
Sarasota	45,817	3.6
Seminole	45,816	1.8
St. Johns	15,219	0.7
St. Lucie	36,254	2.3
County Group 5	62,456	0.8
County Group 6	11,405	2.9
Volusia	65,711	5.2
County Group 7	25,627	6.5
Florida	**2,594,908**	**4.6**

Notes: This summary table uses the 50 percent of the Area Median Income threshold, set at the minimum for county groups—County Group 1: Columbia, Levy, Bradford, Gilchrist, Dixie, and Union; County Group 2: Highlands, DeSoto, Hardee, Okeechobee, Hendry, and Glades; County Group 3: Nassau and Baker; County Group 4: Putnam and St. Johns; County Group 5: Sumter and Lake; County Group 6: Suwannee, Taylor, Madison, Hamilton, and Lafayette; and County Group 7: Walton, Washington, Holmes, and Bay. The full dataset with all household characteristic variables is available on request.
Sources: HUD's Picture of Subsidized Households; American Housing Survey

Exhibit 1 results reveal that far fewer than one in four of Florida's eligible households receives a voucher. Statewide, the calculations show that less than 5 percent of eligible households receive subsidies at the standard eligibility income threshold of 50 percent of the AMI. Moreover, the coverage varies significantly between counties. In several Florida counties, the coverage hovers near 1 percent. Even counties with the highest rates of coverage are below 10 percent. Overall, the statewide voucher gap in Florida is approximately 2.5 million households.

These findings highlight significant disparities in voucher availability across Florida and demonstrate that the actual gaps are much greater than one might believe, given the standardly reported national allocation rate of 25 percent. Hopefully, this methodology can serve as a guide to other fine-grained voucher gap calculations, shed light on the reality of wide voucher gaps, and enable more nuanced discussions of this issue at the local, state, and national levels.

Conclusion

Large voucher gaps represent not only statistics but also families likely struggling to secure affordable housing. The current allocation rates of approximately 5 percent statewide and less than 1 percent in several Florida counties signal a critical failure in meeting the housing needs of Floridians. This situation demands immediate and robust policy interventions at local, state, and national levels. These interventions could include increased federal appropriations for HCVs alongside strategies aimed at enhancing the program's efficacy, such as various landlord participation incentives, efforts to streamline bureaucratic processes, and increased program transparency.

Researchers are encouraged to apply the methodology presented here to calculate county-level voucher gaps across other states or on a national scale to extend this analysis beyond its initial scope. The methodology should be applicable to most other states, requiring only the acquisition of relevant state data, including, for example, county-level inmate counts and crosswalk data to address any missing county information. All data cleaning and calculation scripts are accessible via the GitHub repository, which researchers are invited to clone and modify according to their specific needs (Dabney, 2024).

In addition to the wide voucher gaps that this calculation revealed, an initial data review also showed a significant incongruence between the demographic profiles of households receiving vouchers and those eligible for them. This discrepancy raises questions about possible sociocultural explanations for this result or even systemic issues present in the voucher allocation process. Future research should focus on understanding the root causes of this phenomenon. In addition, examining whether similar patterns exist in other states could be valuable. Such investigations could contribute to a broader understanding of equity issues in voucher allocation and identify targeted interventions to address them.

Policy reforms are urgently needed to meet the demand for affordable housing. Through a combination of research and proactive policy initiatives, it may be possible to bridge the voucher gap and make affordable housing accessible for all who need it.

Acknowledgments

The author would like to express gratitude to the DeVoe L. Moore Center for providing resources and support that made this research possible and, in particular, to Dr. Samuel Staley and Dr. Crystal Taylor for their invaluable insights and guidance throughout the research process.

Special thanks are also extended to research assistants Elizabeth Miller, Eliza Terziev, and Gabriel Valvassori for their dedicated efforts in data collection, research tasks, and valuable input that significantly contributed to the completion of this work.

Author

Shane Dabney is a Ph.D. candidate in philosophy and the Housing Affordability Research Team manager at the DeVoe L. Moore Center at Florida State University.

References

Dabney, Shane. 2024. “sdabney5 / HCVGAPS” GitHub repository. https://github.com/sdabney5/HCVGAPS.

IPUMS. n.d. “HHWT: Household Weight, Technical Variables List, Description.” IPUMS USA. https://usa.ipums.org/usa-action/variables/HHWT#description_section.

Missouri Census Data Center (MCDC). n.d. "Geocorr 2018: Geographic Correspondence Engine." University of Missouri Center for Health Policy. https://mcdc.missouri.edu/applications/geocorr2018.html.

Reina, Vincent, Claudia Aiken, and Jenna Epstein. 2021. *Exploring a Universal Housing Voucher*. The Housing Initiative at Penn. Philadelphia, PA: University of Pennsylvania. https://www.housinginitiative.org/universal-voucher.html.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. 2024. "IPUMS USA: Version 15.0, 2017–2021 5-Year American Community Survey." Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V15.0.

———. 2023. "IPUMS USA: Version 14.0, 2021 5-Year American Community Survey." IPUMS. https://doi.org/10.18128/D010.V14.0.

U.S. Department of Housing and Urban Development (HUD). n.d.a. "Income Limits." https://www.huduser.gov/portal/datasets/il.html#year2021.

———. n.d.b. "Dataset/Assisted Housing: National and Local." https://www.huduser.gov/portal/datasets/assthsg.html.

———. n.d.c. "A Picture of Subsidized Households General Description of the Data and Bibliography." https://www.huduser.gov/portal/datasets/assthsg/statedata98/descript.html.

Additional Reading

DeLuca, Stefanie, Philip M.E. Garboden, and Peter Rosenblatt. 2013. "Segregating Shelter: How Housing Policies Shape the Residential Locations of Low-Income Minority Families," *The ANNALS of the American Academy of Political and Social Science* 647 (1): 268–299.

Ellen, Ingrid Gould. 2015. "Housing Low-Income Households: Lessons from the Sharing Economy?" *Housing Policy Debate* 25 (4): 783–784.

Graves, Erin. 2016. "Rooms for Improvement: A Qualitative Metasynthesis of the Housing Choice Voucher Program," *Housing Policy Debate* 26 (2): 346–361.

Joint Center for Housing Studies (JCHS). 2017. *America's Rental Housing 2017*. Cambridge, MA: Harvard University. https://www.jchs.harvard.edu/sites/default/files/media/imp/harvard_jchs_americas_rental_housing_2017_0.pdf.

Keene, Danya E., Linda Niccolai, Alana Rosenberg, Penelope Schlesinger, and Kim M. Blankenship. 2020. "Rental Assistance and Adult Self-Rated Health," *Journal of Health Care for the Poor and Underserved* 31 (1): 325–339.

Kim, Huiyun. 2022. "Failing the Least Advantaged: An Unintended Consequence of Local Implementation of the Housing Choice Voucher Program," *Housing Policy Debate* 32 (2): 369–385.

Leopold, Josh. 2012. "The Housing Needs of Rental Assistance Applicants," *Cityscape* 14 (2): 275–298.

Sard, Barbara, and Thyria Alvarez-Sanchez. 2011. *Large Majority of Housing Voucher Recipients Work, Are Elderly, or Have Disabilities*. Washington, DC: Center on Budget and Policy Priorities.

Shinn, Marybeth, Scott R. Brown, Brooke E. Spellman, Michelle Wood, Daniel Gubits, and Jill Khadduri. 2017. "Mismatch Between Homeless Families and the Homelessness Service System," *Cityscape* 19 (3): 293–307.

Watson, Nicole Elsasser, Barry L. Steffen, Marge Martin, and David A. Vandenbroucke. 2017. *Worst Case Housing Needs: 2017 Report to Congress*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. https://www.huduser.gov/portal/publications/Worst-Case-Housing-Needs.html.