Comparing Sewage Disposal Method in the American Housing Survey to Administrative Records in Florida

SEHSD Working Paper Number FY 2023-30
September 2023

Patrick Madamba
U.S. Census Bureau[1]

Introduction

This paper assesses the accuracy of respondent-reported sewage disposal methods in the American Housing Survey[2] (AHS) as compared to publicly available administrative records.[3] The study is limited to the state of Florida and focuses on data collected in the AHS from 2015 to 2021. The AHS data and administrative data are classified into two main categories of sewage disposal, namely public sewer systems and private septic systems, which I refer to throughout this paper as "sewer" and "septic," respectively.

I also look at the prevalence of public sewers in Florida according to the AHS, meaning the number of households that reported being connected to a public sewer. These are compared to administrative data to evaluate whether the AHS may be over- or underestimating the prevalence of public sewers. The effect of "do not re-ask" dependency is also examined, wherein AHS housing units that have reported being connected to a public sewer in a previous cycle are assumed to be unlikely to switch to private septic and are no longer asked about their sewage disposal type.

Our two research questions are:

1. To what extent do the self-reported estimates from AHS differ from estimates based on the administrative data?

[1] This paper is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Disclosure Review Board (DRB) approval number: CBDRB-FY23-POP001-0109).

I would like to thank Jeremy Engelberg for his work on the internal analysis comparing AHS sewage disposal estimates to admin data in Maryland, which this paper builds on. I also appreciate the various comments from staff at the U.S. Census Bureau and the U.S. Department of Housing and Urban Development.

[2] More information on confidentiality protection, methodology, sampling and nonsampling error, and definitions is available at <https://www.census.gov/programs-surveys/ahs/>.

[3] All comparative statements in this paper have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 10 percent significance level.

2. To what extent are AHS estimates affected by the use of dependent interviewing for sewage disposal type?

Methodology

Florida was chosen as the focus of this study because the AHS sample is consistently representative enough to make state-level estimates, and administrative records for sewage are publicly available for Florida. The AHS data used in this analysis were collected every odd-numbered year from 2015 to 2021. The administrative data (also referred to as "admin data" or "admin records") were downloaded from the Florida Department of Health website from its Florida Water Management Inventory (FLWMI) project,[4] which gathers information about drinking water source and wastewater treatment method from multiple stakeholders[5] across the state. The project website notes that the initial inventory was completed in October 2016, but it is continually updated to address gaps in parcel data where public water and wastewater disposal providers did not submit data. The data used in this analysis were downloaded from the Florida Department of Health website in August 2022. The years when wastewater disposal method was updated for different parcel units range from as far back as 2009 until as recently as 2022. Data are available in the form of ESRI® shapefiles,[6] which cover all land parcels in all 67 counties of Florida. Attached to each parcel are several variables, including geographic information, method by which drinking water is delivered, method by which domestic wastewater is disposed, and the source of the data.

Sewage disposal in the 2015-2021 AHS data (variable name *SEWTYPE*) has 10 possible values with categories including public sewer, some type of septic tank or cesspool,[7] chemical toilet, outhouse, other type of sewer system, or no sewer system. To match these categories with the admin data, *SEWTYPE* was recoded into three categories. Public sewer responses were categorized as "Sewer." Households that reported having a septic tank or cesspool, chemical toilet, outhouse, or other sewer type were all considered to have a private system and were categorized as "Septic." Those that had either reported no sewer system or did not report a response were classified as "N/A" and were excluded from the analysis.

[4] An overview of the project and a link to the publicly available data is available on the FLWMI website.

[5] More information about the current project status and how data was collected is available on the details page of the FLWMI website.

[6] A shapefile stores nontopological geometry and attribute information for the spatial features in a data set. The geometry for a feature is stored as a shape comprising a set of vector coordinates. The ESRI® website has detailed technical documentation about shapefiles.

[7] There are five different categories for type of septic tank or cesspool, namely: a) standard septic tank and subsurface leach field; b) uses a pump to distribute wastewater; c) elevated above natural soil surface; d) applies treated wastewater; and, e) other type.

In the Florida administrative records, sewage disposal has nine possible values. These categories are listed in the first column of Table 1 below. According to the FLWMI data dictionary,[8] the wastewater disposal method for each parcel is "determined by evaluating all of the wastewater values assigned to the parcel from the various sources that provide wastewater information" to the FLWMI. Some of the data sources include parcel data from the Florida Department of Revenue, construction and operating permits for wastewater systems from the Florida Department of Health, and facility information from individual utilities and treatment plants. Because there are multiple data sources, the inventory assigns the qualifiers "known," "likely," and "somewhat likely" based on relative confidence in the information available for each parcel. For instance, parcels with information based on utility account records would receive the qualifier "known," while those based on septic permits that were never inspected may receive the qualifier "likely." For this paper, I combine these multiple likelihood categories into simply "Sewer" and "Septic" to make the admin data comparable to the recoded AHS data. Table 1 below details how the admin data categories were recoded for this paper.

Table 1. Summary of Recodes for Wastewater Disposal Method Variable on FL Administrative Records.

Administrative data categories *(**Code**: Meaning)*	**Recoded categories** *(**Code**: Meaning)*
KnownSewer: Known central sewer **LikelySewer**: Likely central sewer **SWLSewer**: Somewhat likely central sewer	**1**: Sewer
KnownSeptic: Known onsite septic system **LikelySeptic**: Likely onsite septic system **SWLSeptic**: Somewhat likely onsite septic system	**2**: Septic
NA: N/A, not built	**3**: N/A, Not Built
UNK: Unknown, no data	**4**: Unknown, No Data
UNDT: Undetermined, conflicting data	**5**: Conflicting Data

In the admin data, the category "Not Built" means that there is no requirement for domestic wastewater disposal on the parcel and thus none was built. The "No Data" category means that no information was available for wastewater disposal method, or the requirement for wastewater disposal could not be determined. The "Conflicting Data" category implies that two or more data sources had equal opposing values for domestic wastewater disposal, and therefore the correct disposal method for that parcel could not be determined. An example of conflicting data would be if data from the utility company shows that a parcel is connected to a sewer line, but the Environmental Health Database also indicates that a permit for septic has been issued for that same parcel. Similar to the "N/A" responses in the AHS data, the "Not Built," "No Data," and "Conflicting Data" categories were dropped from the analysis. The Limitations section of this paper explores why the admin data may have had these codes for certain parcel data that were joined to AHS sample.

[8] The FLWMI GIS Data Dictionary outlines the contents of the final GIS datasets used in the project.

Matching Process

Wastewater information from the administrative data was matched to individual AHS sample records using the spatial join tool (Arc Toolbox > Overlay > Spatial Join) in ArcMap™. AHS sample records were mapped as points using latitude and longitude (which I also refer to as "lat" and "long," respectively). The land parcels in the admin data were mapped as polygons and attached to each AHS sample housing unit using the corresponding parcel polygon of each lat/long point. Each lat/long point that did not fall inside a polygon (e.g., the point was mapped onto the middle of a road) was joined to the parcel polygon closest in distance to it. Figure 1 below illustrates how the spatial join was conducted.

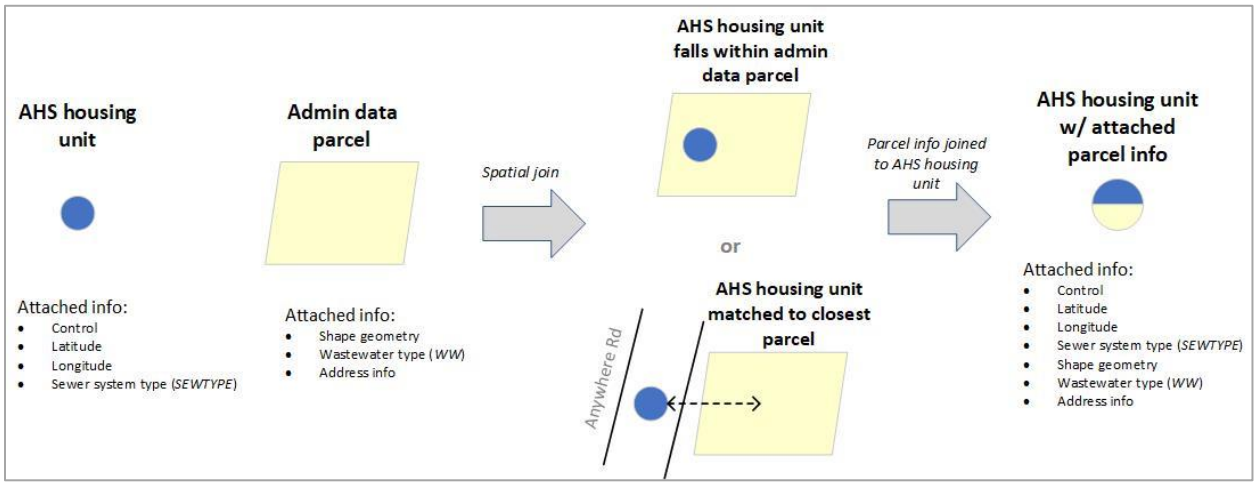

Figure 1. Illustration of Spatial Join Process to Attach Administrative Records to AHS Sample.

Table 2 below summarizes how many housing units fell within the bounds of an admin data parcel and how many had to be matched to the nearest parcel. Across all years, most housing units were contained within an admin data parcel. The Census Bureau made improvements to the accuracy of its geographic coordinates in the AHS sample in 2019, which may account for the increase in AHS sample that were mapped within a parcel.

Table 2. Distribution of AHS Sample Based on Spatial Join Method Used.

Spatial join method	2015		2017		2019		2021	
	Freq	%	Freq	%	Freq	%	Freq	%
0: AHS sample contained in admin data parcel	4,200	87.3	4,200	87.3	4700	95.5	4500	95.4
1: AHS sample matched to nearest admin data parcel	600	12.7	600	12.7	200	4.5	200	4.6

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Due to the size of each county-level shapefile, merging all 67 counties into one statewide shapefile was not possible. Instead, I attached parcel information to the AHS sample records one county at a time, and then stacked all counties back together. To streamline the process, the FL counties where AHS sample are present were identified using the *COUNTY_2010* variable in the AHS, which contains the Federal

Information Processing Standards (FIPS) three-digit county codes for 2010. The spatial join was then done for only those counties, about 34 counties per survey year.

Match Quality

I evaluated the quality of the point-to-parcel matching process by comparing physical addresses from the admin data with the basic street address of each AHS housing unit. Extraneous blanks were removed from the admin data's physical address, which included both housing number and street name, and the resulting variable was named "*ADMINGEO*." From the AHS sample, housing number and street name were similarly cleaned and concatenated, with the resulting variable named "*AHSGEO*." Using SAS® software, I compared *ADMINGEO* to *AHSGEO* and gave it a match quality of 'Match – exact, automated' if they were equal. Those that did not match exactly were exported to a spreadsheet and manually evaluated. The table below summarizes how match quality was quantified.

Table 3. Summary of Match Quality Assessment for Each Housing Unit.

Value given	Meaning	Description	Examples	
			ADMINGEO	***AHSGEO***
0	Not a match	*ADMINGEO* completely different from *AHSGEO*.	479 Somewhere Cir	123 Anywhere Ln
1	Match – close enough	*ADMINGEO* and *AHSGEO* had the same street name, but the house number was not an exact match.	1) 678 Someplace Dr 2) 3rd St	1) 670 Someplace Dr 2) 1052 3rd St
2	Match – exact, manually evaluated	*ADMINGEO* did not word-for-word match *AHSGEO* but was determined to be the same address.	1) 123 West Ave 2) 451 Anywhere 3) 17 Somewhere Lp 4) 15 USHY 5) 19 St	1) 123 Ave W 2) 451 Anywhere Rd 3) 17 Somewhere Loop 4) 15 US Highway 5) 19th St
3	Match – exact, automated	Using SAS® software, *ADMINGEO* = *AHSGEO*.	123 Anywhere Ln	123 Anywhere Ln
M	Missing address information	The attached administrative data did not include address information.		123 Anywhere Ln

Table 4 below summarizes the address match quality across all years. For all years, the percentage of addresses that did not match at all ranged from 6.8% to 11.5%. If we consider "close enough" as sufficient to constitute a match, the overall address match rates are between 68.3% and 73.2%. The noticeable drop in non-matches (and increase in overall match rates) from 2017 to 2019 could be due to the improvements in geographic coordinates for the AHS sample applied in 2019. It is also important to note that around 20% of housing units had missing address information ('M') in the admin data across all years, and if this information had been available in the admin data it is possible that the overall address match rate might in fact have been higher than is reported here. Other limitations to the address match rates include any human error in the manual evaluation and the address information in the admin data

possibly being outdated. It is important to emphasize that this analysis was only done to roughly measure the accuracy of the point-to-parcel match in ArcMap™; no housing units were removed from the subsequent analyses based on their address match quality being 'Not a match' or 'Missing address information.'

Table 4. Summary of Address Match Quality for Florida 2015-2021.

	2015		2017		2019		2021	
Address match quality	**Frequency**	**Percent***	**Frequency**	**Percent***	**Frequency**	**Percent***	**Frequency**	**Percent***
0: Not a match	500	10.8	550	11.5	350	6.8	400	8.6
1: Match - Close enough	700	13.9	650	14.0	650	13.2	600	13.3
2: Match - Exact, Manually evaluated	1,000	21.6	1,000	21.8	1,200	24.3	1,200	26.5
3: Match - Exact, Automated	1,600	33.4	1,600	32.5	1,800	35.7	1,500	32.4
M: Missing address information	1,000	20.4	950	20.3	1,000	20.0	900	19.1
Overall address match rate (sum of categories 1-3)	68.9%		68.3%		73.2%		72.3%	

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Results and Discussion

This paper examines four groups: all housing units, owner-occupied housing units, renter-occupied housing units, and vacant/URE (usual residence elsewhere) housing units. Three types of analysis were conducted for each group. The first looked at sewage disposal agreement rates, comparing the self-reported AHS response to the value in the admin data. The second analysis looked at prevalence rates of public sewer, meaning the percentage of housing units in the AHS that reported sewer compared to the percentage of housing units whose admin data indicated sewer. This was to determine if the AHS is over- or underestimating the overall prevalence of public sewers compared to the admin data for Florida. The third analysis examined the percentage of mismatches within the group whose admin data indicated septic but whose AHS value was sewer. Towards the end of this section, I also look at the effect of "do not re-ask" (DNR) dependency on sewage disposal estimates.

It is important to clarify the records from which Florida estimates in this section are produced. As mentioned in the Methodology section, AHS sample that were classified as 'N/A' (AHS self-report was no sewer system or don't know or refused) were dropped from this analysis. In addition, AHS records that were matched to an admin data parcel whose wastewater disposal method was "Not built," "No data," or "Conflicting data" were also dropped from this analysis. To summarize, I limited this analysis to the subset of housing units with an observed value in both the AHS and admin data. Table 5 shows the unweighted numbers of housing units for which the AHS collected sewage information, housing units dropped, and the remaining housing units from which weighted estimates were subsequently produced. All percentages presented in the results are weighted estimates from the remaining sample produced using AHS replicate weights.

Table 5. Unweighted Numbers of AHS Housing Units in Florida That Were Dropped from the Analysis Due to a Not Applicable Sewage Disposal Response in the AHS or Administrative Data.

AHS Survey Year	Initial number of housing units in AHS Florida sample	Number of dropped housing units with not applicable sewage disposal response	Percent of initial housing units dropped	Number of remaining housing units from which estimates are produced for this analysis
2015	4,900	700	14.5%	4,200
2017	4,900	650	13.6%	4,100
2019	4,900	650	12.8%	4,300
2021	4,700	650	13.6%	4,000

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

All Housing Units

Table 6 summarizes the sewage disposal agreement rates in Florida from 2015 to 2021 based on the agreement matrices in Tables B-1 to B-4 in Appendix B. Three percentages are presented: the

percentage out of the total for which sewage disposal method in both the AHS and admin data is sewer, the percentage out of the total for which both is septic, and the combined agreement rate which is the sum of the sewer-sewer and septic-septic agreement. The combined agreement rate is at around 85% for all years in Florida. Over this observed period, the estimated sewer-sewer agreement rate and septic-septic agreement rate have remained statistically unchanged.

Table 6. Sewage Disposal Method Agreement Rates Between AHS and Admin Data for All Housing Units.

AHS Survey Year	Sewer-sewer agreement out of the total (%)	Septic-Septic agreement out of the total (%)	Combined agreement rate for both sewer and septic (%)
2015	71.5	13.8	85.3
2017	73.5	12.1	85.6
2019	74.7	10.8	85.6
2021	75.1	9.7	84.8

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

The next analysis compares the prevalence of public sewers reported in the AHS to their prevalence rate in the admin records. This may provide some insight into whether the AHS is over- or underestimating sewers (and conversely, septic). Table 7 shows that the prevalence rate of sewers in the AHS is higher than the prevalence rate in admin records across all years. This suggests that the AHS may be overestimating the prevalence of public sewers. There is a significant increase in the AHS prevalence rate from 83.2% (SE 1.9%) in 2015 to 89.4% (SE 2.0%) in 2021, but the estimated admin data prevalence rate remained statistically unchanged over the same period.

Table 7. Prevalence Rates of Public Sewer as Reported in the AHS and in Administrative Data for All Housing Units.

AHS Survey Year	Total public sewer – AHS reported (%)	Total public sewer – Admin data (%)	Difference (AHS vs Admin data)
2015	83.2	74.4	8.9
2017	86.7	74.7	12.0
2019	88.3	75.6	12.7
2021	89.4	76.0	13.4

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Appendix tables B-1 to B-4 present agreement matrices that report the frequencies for how often the AHS and admin data show conflicting responses. The agreement matrices show that the most common type of mismatch is housing units that self-reported sewer on the AHS but have septic according to the admin records. The converse of this, septic in the AHS but sewer in the admin data, occurs to a lesser extent. For example, the estimates in Table B-4 show that roughly 6.6 million Florida housing units in 2021 were connected to sewer based on the admin data – and that only about 77 thousand or 1.2% of these

units self-reported septic in AHS. In contrast, among the 2.1 million housing units estimated to have septic based on the admin data, about 1.2 million or nearly 60% of these units self-reported sewer in the AHS.

Figure 2 focuses on the set of housing units whose admin data indicates septic, presenting the percentage of these housing units for which the AHS reported sewer. The results show that 46% (SE 4.9%) self-reported as sewer on the 2015 AHS, and that mismatch rate has increased over time to 59.6% (SE 6.1%) in the 2021 AHS.

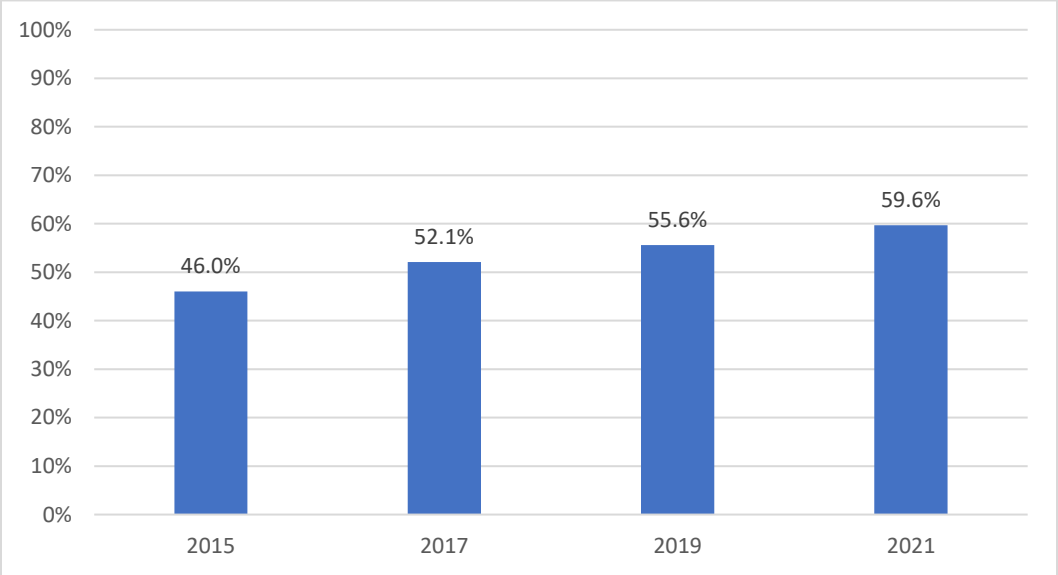

Figure 2. Percentage Mismatch with AHS Self-Reporting Sewer out of the Group of Housing Units Whose Administrative Data Indicate Septic, for All Housing Units (Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records).

Owner-Occupied and Renter-Occupied Housing Units

The groups of owner-occupied and renter-occupied housing units were analyzed side-by-side to see if tenure affects how AHS self-reported sewage disposal type agrees with admin records. Table 8 below summarizes agreement rates from matrices presented in Appendix C for owner-occupied units and Appendix D for renter-occupied units. The combined agreement rates for all years, for both owners and renters, appear to be close, but by separating the estimated percentages out of total for both sewer-sewer agreement and septic-septic agreement we start to see some differences. Renters show higher sewer-sewer agreement percentages out of total than owners, but much lower septic-septic agreement, which notably decreased from 8.8% (SE 1.5%) in 2015 to 3.5% (SE 1.4%) in 2021. It seems that compared to owners, renters are more likely to self-report sewer when the administrative records also indicate the parcel is sewer but are less likely to report septic when the administrative records indicate the parcel is septic.

Table 8. Sewage Disposal Method Agreement Rates Between AHS and Administrative Data for Owner-Occupied and Renter-Occupied Housing Units.

AHS Survey Year	Sewer-sewer agreement out of the total (%)		Septic-septic agreement out of the total (%)		Combined agreement rate for both sewer and septic (%)*	
	Owner	Renter	Owner	Renter	Owner	Renter
2015	68.1	75.0	18.2	8.8	86.3	83.8
2017	68.8	79.5	17.1	6.4	85.9	85.9
2019	70.4	79.5	15.5	5.0	85.8	84.6
2021	70.7	81.9	13.8	3.5	84.4	85.4

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

As Table 9 shows, the prevalence of public sewer in the AHS (columns 2 and 5) is higher than the prevalence in the admin data (columns 3 and 6) for all years, regardless of tenure type. This continues to suggest that the AHS may be overestimating the prevalence of public sewers compared to the admin data. However, the differences between AHS and admin data prevalence rates for owners (column 4) are not statistically different from those for renters (column 7).

Table 9. Prevalence of Public Sewer in the AHS and Administrative Data for Owner-Occupied and Renter-Occupied Housing Units.

AHS Survey Year	Owner-occupied			Renter-occupied		
	Total public sewer reported in AHS (%)	Total public sewer in Admin data (%)	Difference (AHS-Admin data)	Total public sewer reported in AHS (%)	Total public sewer in Admin data (%)	Difference (AHS-Admin data)
2015	78.3	71.6	6.7	88.1	78.2	9.9
2017	81.2	70.5	10.7	92.9	80.1	12.8
2019	83.3	71.6	11.7	94.4	80.1	14.4
2021	84.9	71.9	13.0	96.0	82.4	13.7

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Figure 3 focuses on the group of housing units for which the admin data indicates septic, showing the percentage of these owner-occupied and renter-occupied housing units for which AHS respondents self-report sewer. Over time, the percentage of owners with this mismatch increased from 36.0% (SE 4.5%) in 2015 to 50.9% (SE 5.7%) in 2021, while the percentage of renters with this mismatch increased from 59.8% (SE 5.5%) to 80.2% (SE 7.4%) over the same period. This shows that renters are more likely than owners to self-report sewer when their sewage disposal method is indicated in the admin data as septic.

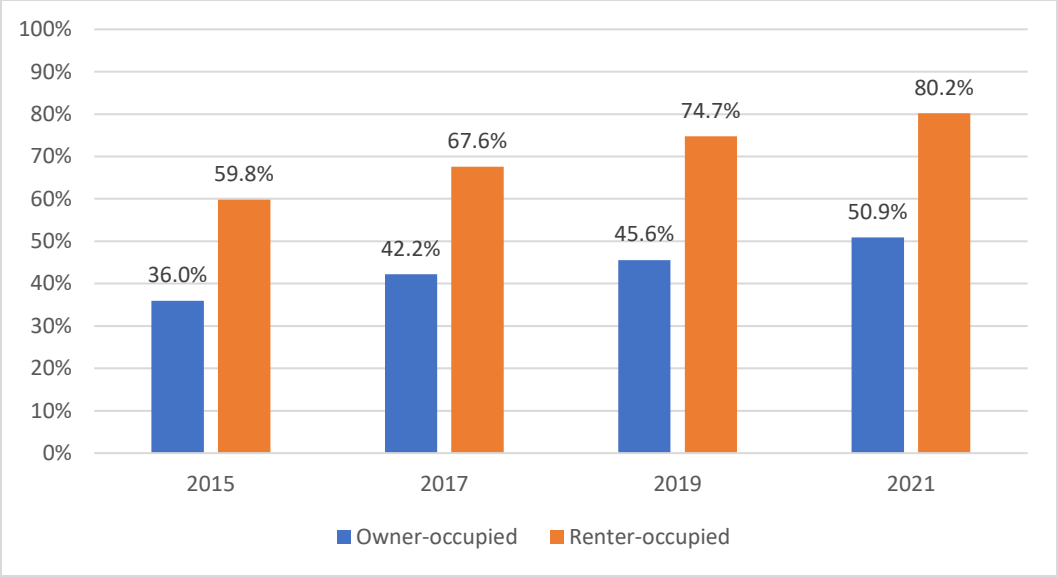

Figure 3. Percentage Mismatch with AHS Self-Reporting Sewer out of the Group of Housing Units Whose Administrative Data Indicate Septic, for Owner-Occupied and Renter-Occupied Housing Units (Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records).

Vacant/URE Housing Units

For vacant and usual residence elsewhere (URE) units, the combined sewage disposal method agreement rate ranges from 84.2% to 86.3% across all years. Over this period, the estimated percentages out of total for both sewer-sewer agreement and septic-septic agreement have remained statistically unchanged.

Table 10. Sewage Disposal Method Agreement Rates Between AHS and Administrative Data for Vacant/URE Housing Units.

AHS Survey Year	Sewer-sewer agreement out of the total (%)	Septic-Septic agreement out of the total (%)	Combined agreement rate for both sewer and septic (%)
2015	74.8	10.3	85.1
2017	77.7	6.6	84.2
2019	80.3	6.0	86.3
2021	79.5	5.7	85.2

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

The vacant and URE units group shows higher sewer prevalence rates in the AHS than in the admin data for all years. However, despite the large magnitude of differences between AHS and admin data

prevalence rates in column 4 of Table 11 below, these are not statistically different from those of the renter-occupied group.

Table 11. Prevalence Rates of Public Sewer as Reported in the AHS and Administrative Data for Vacant/URE Housing Units.

AHS Survey Year	Total Sewer – AHS reported (%)	Total sewer – Admin data (%)	Difference (AHS-Admin data)
2015	88.7	75.8	12.9
2017	92.7	78.4	14.3
2019	93.4	80.8	12.6
2021	94.2	79.6	14.6

Source: U.S. Census Bureau, American Housing Survey. 2015-2021; Florida Water Management Inventory administrative records.

Figure 4 shows the mismatch percentages for the group of vacant/URE housing units whose admin data indicates septic, but whose self-reported value is sewer. Unlike the previous groups examined, the estimated mismatch rate remained statistically unchanged over this period. The vacant/URE mismatch rates are higher than those for owner-occupied units but are not statistically different from those for renter-occupied units.

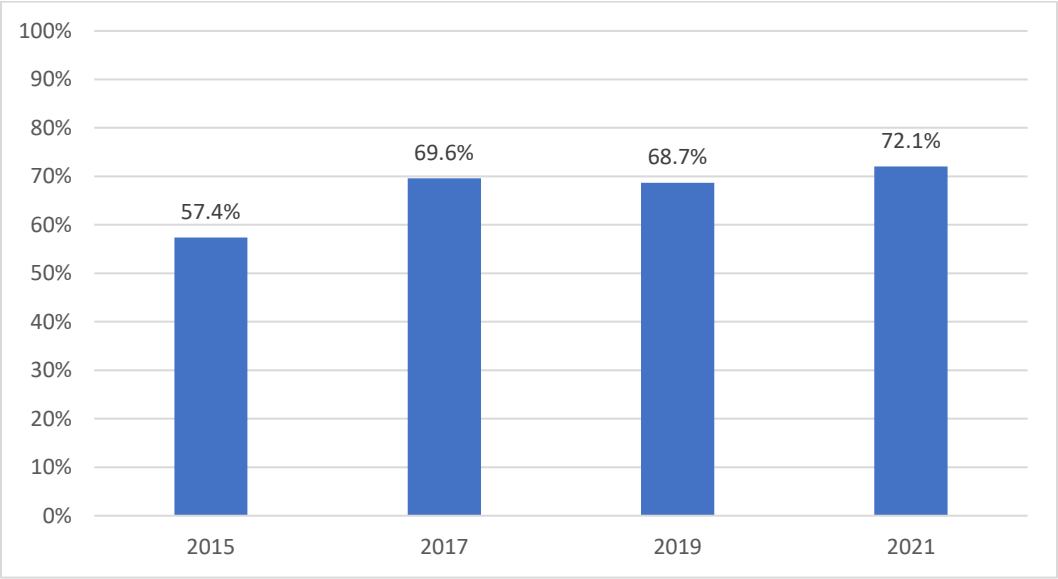

Figure 4. Percentage Mismatch with AHS Self-Reporting Sewer out of the Group of Housing Units Whose Administrative Data Indicate Septic, for Vacant/URE Housing Units (Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records).

Effect of Do Not Re-ask (DNR) Dependency in the AHS Questionnaire

The AHS is conducted using a Computer Assisted Personal Interviewing (CAPI) laptop instrument. The instrument allows for "dependent interviewing," meaning responses to questions from previous surveys can be used in subsequent years to reduce respondent burden.[9] When a respondent reports that the housing unit is connected to public sewer, the CAPI instrument will skip that question in future survey years. This is referred to as "do not re-ask" (DNR) dependency. This approach assumes that housing units connected to a public sewer are unlikely to switch to a private sewage disposal system. The DNR dependency for public sewer began in 2017.

Table 12 shows the extent to which the sample of DNR housing units in each AHS survey year agrees with the sewage disposal method listed in the administrative data over time. These are AHS housing units that were no longer asked their sewage disposal method because they are assumed to still be connected to public sewer. In 2017, the first year of DNR dependency, 13.4% (SE 1.4%) of DNR housing units self-reported sewer but had septic according to the admin data. Most recently, in the 2021 AHS, this percentage was 15.1% (SE 1.1%) of DNR housing units. It is important to note that the estimated mismatch rate remained statistically unchanged over this period.

Table 12. Sewage Disposal Method Agreement Rates Between the AHS DNR Sample and the Administrative Data, for All Housing Units in Florida.

			Admin Data	
		AHS Year	**Sewer (%)**	**Septic (%)**
AHS Data	**Sewer** (DNR housing units on the AHS that reported being connected to public sewer at last interview)	2017	86.6	13.4
		2019	85.3	14.7
		2021	84.9	15.1

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Note: The 2015 AHS was the first wave of the panel and thus does not have DNR dependency. Therefore, only 2017-2021 have DNR dependency and are shown in this table.

Table 13 replicates this analysis for DNR AHS housing units based on tenure. In 2017, 13.1% (SE 1.6%) of owner-occupied DNR units had their sewage disposal type recorded as sewer due to the DNR dependency but had septic in the admin data. The percentage of renters with that mismatch that year was 13.8% (SE 1.7%). In the 2021 AHS, 15.3% (SE 1.5%) of owners had this mismatch, while 14.7% (SE 1.3%) of renters had this mismatch. However, there are no statistical differences in the mismatch rates between owner-occupied units and renter-occupied units. The estimated mismatch rates also remain statistically unchanged over time for both tenure groups.

[9] More information can be found on the AHS documentation for Dependent Interviewing: 2015 and Beyond.

Table 13. Sewage Disposal Method Agreement Rates Between the AHS DNR sample and the Administrative Data, for Owner-Occupied and Renter-Occupied Housing Units in Florida.

			Admin Data				
			Owner-occupied		Renter-occupied		Difference between renter septic vs owner septic
		AHS Year	Sewer (%)	Septic (%)	Sewer (%)	Septic (%)	
AHS Data	**Sewer** (DNR housing units on the AHS that reported being connected to public sewer at last interview)	2017	86.9	13.1	86.2	13.8	0.7
		2019	85.6	14.4	84.9	15.1	0.7
		2021	84.7	15.3	85.3	14.7	0.6

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Note: The 2015 AHS was the first wave of the panel and thus does not have DNR dependency. Therefore, only 2017-2021 have DNR dependency and are shown in this table.

Even though the estimated mismatch rates are unchanged over time and are not statistically different between renters and owners, the fact that the estimated mismatch rates are statistically different from zero suggests that DNR dependency is contributing to an overestimate of housing units connected to public sewer when the administrative records show the parcel as septic. DNR dependency also only allows for septic to sewer transitions, but not the reverse. This drift could be evidenced by percentages of AHS estimates of sewer increasing over time, which we observed earlier in Table 7 with the prevalence rates of AHS sewer increasing from 83.2% (SE 1.9%) in 2015 to 89.4% (SE 2.0%) in 2021. Figure 2 also supports this notion, as we see that of the sample of housing units for which the admin data indicates septic, 46% (SE 4.9%) self-reported the unit as sewer on the 2015 AHS. This mismatch percentage has increased to 59.6% (SE 6.1%) in the most recent 2021 AHS. Further evidence of a drift to less septic would be a decrease in the septic-septic agreement percentage out of total over time, which is observed in Table 8, particularly for renter-occupied housing units whose septic-septic agreement percentage out of total decreased from 8.8% (SE 1.5%) in 2015 to 3.5% (SE 1.4%) in 2021. It seems unlikely that the septic-septic agreement percentage out of total for renters would decrease over time entirely due to renters increasingly not knowing what sewage disposal method their unit uses, or due to more than half of rental units with septic converting to sewer and the AHS not capturing that shift. Instead, it seems more plausible that the drift toward less septic is driven by an accumulation over the years of DNR housing units initially recording their sewage disposal method as sewer and that being carried forward into subsequent survey years.

To further understand the impact of dependency, the 2023 AHS will re-ask all housing units the public sewer question.

Limitations

Not Applicable Values in Admin Data

The Methodology section discussed how the admin data included some "Not applicable" response codes: "Not built," "No data," and "Conflicting data." These housing units were removed from the sample that was subsequently analyzed. Here I attempt to evaluate its effect and explain why these categories are excluded.

A summary of the total weighted percentages of these "Not applicable" responses from the admin data is shown below:

Table 14. Sum of "Not Applicable" Categories in Administrative Data.

AHS Year	Percentage Not built + No data + Conflicting data (%)
2015	12.0
2017	12.5
2019	11.7
2021	13.1

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

The Florida Water Management Inventory (FLWMI) project website notes that the initial inventory was completed in October 2016 with source data as old as 2009 but is continually updated from parcel information provided to the project team by public and private entities. It is possible that these "Not applicable" categories are housing units for which the admin data does not have updated data and the AHS has more recent data. Table 15 shows the unweighted percentage of housing units with "Not applicable" categories that were last updated in the admin data before the AHS survey year and therefore may not be up to date compared to AHS data.

Table 15. Summary of Year That Wastewater Disposal Method Was Updated in the Administrative Data for Sample with Not Applicable Categories.

AHS Year	Year that domestic wastewater disposal was updated in admin data
2015	6.2% were last updated in or before 2014. Remaining were last updated between 2015-2022.
2017	66.4% were last updated in or before 2016. Remaining were last updated between 2017-2022.
2019	92.2% were last updated in or before 2018. Remaining were last updated between 2019-2022.
2021	91.9% were last updated in or before 2020. Remaining were last updated between 2021-2022.

Source: Florida Water Management Inventory administrative records.

Large percentages of these housing units had their admin records last updated prior to each AHS year, particularly in 2019 and 2021. It is possible that at the time the admin records were collected, there was indeed no wastewater disposal built, or no data was available, or there was conflicting data, and that the AHS has more recent data than the admin records. This is one limitation of the admin data used and is why the "Not applicable" categories were removed from the subsequent analysis.

Other States and Future Research

Extending this analysis to other states is contingent on two prerequisites. The first is that the AHS should have a sample large enough that is representative for the entire state. The second is that administrative records on the sewage disposal type for all parcels in the state should be publicly available. An internal analysis was previously conducted for Maryland, for which the AHS has state-level estimates and parcel data is publicly available from the Maryland Department of Planning. The internal analysis found that the AHS self-reported values and administrative records had a high agreement rate, but there was evidence to suggest that renters may be less likely than other groups to report their home being connected to septic when the admin records also indicate septic. The Maryland analysis also found that dependent interviewing did not have a significant effect on the agreement rates. One limitation of the analysis, however, was that it only examined 2017 data, and therefore could not measure changes over time. Another limitation was that while the sample was representative for the state, Maryland still had a smaller sample size, which may have caused groups such as the vacant/URE group to not show any significant differences. These limitations are addressed in this paper, with Florida having a larger sample size and data from 2015 to 2021 being analyzed.

For future studies, it would be beneficial to determine if there is a geographical component to the results in this paper by applying the methodology to other states. Ideally, the comparison states would have sample sizes comparable to or larger than Florida, such as New York, Texas, or California. These states are also in different Census Divisions than Maryland and Florida. However, the feasibility of these future studies still depends on administrative records being publicly available in those states. In addition, with dependent interviewing being removed from the public sewer question in the 2023 AHS, it could be worthwhile to return later to Florida to see if there are any significant changes to agreement rates.

Conclusions

The first research question seeks to determine the extent to which AHS estimates differ from those on available administrative data. To evaluate this, three types of analysis were conducted. The first looked at agreement rates between the AHS self-reported value and the administrative records value when both are sewer (sewer-sewer) and both are septic (septic-septic). The second analysis compared the prevalence rates of public sewer, defined as the total percentage of weighted housing units that reported sewer, on the AHS versus on the admin records. The third analysis took the subset for which the admin data recorded the housing unit as septic, and from that calculated the mismatch percentage of units whose self-reported value was sewer. These three analyses were done for four groups: all housing units, owner-occupied housing units, renter-occupied housing units, and vacant or usual residence elsewhere (URE) housing units.

The combined sewage disposal agreement rates (sewer-sewer agreement + septic-septic agreement) appear to be stable at around 85% for all years in Florida, regardless of tenure or interview status. However, separating out the agreement percentages out of total tells a different story, particularly for renter-occupied units. The sewer-sewer agreement pecentages out of total appear stable over the period 2015-2021 for all groups, except for renters where it increased from 75.0% (SE 1.3%) to 81.9% (SE 1.3%). Meanwhile, the septic-septic agreement percentage out of total decreased over time only for renters, from 8.8% (SE 1.5%) in 2015 to 3.5% (SE 1.4%) in 2021.

Comparing prevalence rates of public sewers suggests that the AHS tends to overestimate public sewers compared to the administrative records. For all housing units, most recently in 2021, the sewer prevalence rate in the AHS was 89.4% (SE 2.0%), which was significantly higher than that of the admin data at 76.0% (SE 1.8%), a difference of about 13.4 percentage points. The owner-occupied, renter-occupied, and vacant/URE groups all showed higher sewer prevalence rates as well in the AHS compared to the admin data, across all years. The prevalence rate of public sewer as self-reported in the AHS is also increasing over time for all groups except the vacant/URE group. For instance, the AHS prevalence rate for owners increased from 78.3% (SE 2.1%) in 2015 to 84.9% (SE 2.4%) in 2021, while the admin data prevalence rate for owners remained stable over time, at 71.6% (SE 2.6%) in 2015 and at 72.0% (SE 2.5%) in 2021. Renters were the only group to show a significant increase in both AHS and admin data prevalence rates; from 2015 to 2021, the AHS prevalence rate for renters increased from 88.1% (SE 1.6%) to 96.0% (SE 1.4%), while the admin data prevalence rate increased from 78.2% (SE 1.5%) to 82.4% (SE 1.3%).

The most common type of mismatch is housing units that self-report sewer on the AHS but whose admin records indicate septic. Focusing on the set of housing units whose admin records indicate septic, the mismatch rates range from 46.0% (SE 4.9%) to 59.6% (SE 6.1%) for all housing units. Renter-occupied units are again a notable group, with mismatch rates increasing from 59.8% (SE 5.5%) in 2015 to as high as 80.2% (SE 7.4%) in 2021. These are much higher than the corresponding mismatch rates in the owner-occupied group, which range from 36.0% (SE 4.5%) to 50.9% (SE 5.7%).

Returning to the research questions, looking holistically at agreement rates, prevalence rates, and mismatch rates, AHS estimates do appear to differ from those on admin data in Florida, particularly for renter-occupied housing units. The tenure-related variances could be due to renters not knowing the

sewage disposal method used by their rental unit compared to owners. These differences could also be related to dependent interviewing having a larger effect.

This leads in to the second research question, which focuses on the extent of the effect of dependent interviewing. The percentage of housing units that recorded their sewage disposal method as public sewer due to do not re-ask (DNR) dependency but have septic in the admin records was 13.4% (SE 1.4%) in 2017, the first year when DNR was applied to the AHS sample. Most recently in 2021, this mismatch for DNR unit was 15.1% (SE 1.1%). This mismatch percentage was stable over the years investigated, with no statistically significant differences between renters and owners.

To summarize, there are three main points from comparing AHS estimates with admin data. First, the groups for all housing units, owners, and renters show increasing prevalence rates of sewer in the AHS. Second, the mismatch percentages for units with septic in the admin data but sewer in the AHS are increasing over time for all groups, except the vacant/URE group. Third, renters show decreasing septic-septic agreement percentages out of total over time. Considering the evidence from these three points together with the DNR analysis, it appears that dependent interviewing may be causing AHS estimates to drift toward less septic and more sewer over time, particularly for renter-occupied units.

Appendix A

As mentioned in the Methodology section, the wastewater disposal method for each parcel in the administrative data is determined by evaluating several data sources that provide information to the FLWMI. Because there are multiple data sources, the inventory assigns the qualifiers "known," "likely," and "somewhat likely" based on relative confidence in the information available for each parcel. For the purposes of this paper, I simplified the assumption by ignoring the qualifiers and classifying parcels as either "sewer" or "septic" if an applicable value was available.

Table A-1 below presents these agreement rates out of the sample for which the AHS value is either sewer (for the known/likely/somewhat likely sewer categories) or septic (for the known/likely/somewhat likely septic categories) broken down by these likelihood categories. This is shown to explore the likelihood that an admin data parcel is connected to a particular sewage disposal system has any variation on agreement rates with the AHS self-reported value in 2021.

Table A-1. Agreement Rates Between AHS and Administrative Data Broken Down by Likelihood Category for 2021 in Florida.

		AHS Agreement %
Admin Records	Known sewer	99.1
	Likely sewer	96.4
	Somewhat likely sewer	98.0
	Known septic	49.7
	Likely septic	34.8
	Somewhat likely septic	43.7

Source: U.S. Census Bureau, American Housing Survey, 2015-2021; Florida Water Management Inventory administrative records.

Appendix B

Table B-1. AHS and Administrative Data for 2015 for All Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	5,716,000	71.5	943,000	11.8	6,659,000	83.2
	Septic	234,000	2.9	1,107,000	13.8	1,341,000	16.8
	Total	5,950,000	74.4	2,050,000	25.6	8,000,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2015; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table B-2. AHS and Administrative Data for 2017 for All Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	6,057,000	73.5	1,088,000	13.2	7,145,000	86.6
	Septic	101,000	1.2	1,000,000	12.1	1,101,000	13.4
	Total	6,159,000	74.7	2,088,000	25.3	8,246,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2017; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table B-3. AHS and Administrative Data for 2019 for All Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	6,368,000	74.7	1,155,000	13.5	7,522,000	88.3
	Septic	77,000	0.9	922,000	10.8	999,000	11.7
	Total	6,444,000	75.6	2,077,000	24.4	8,521,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2019; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table B-4. AHS and Administrative Data for 2021 for All Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	6,551,000	75.1	1,248,000	14.3	7,799,000	89.4
	Septic	77,000	0.9	844,000	9.7	921,000	10.6
	Total	6,628,000	76.0	2,092,000	24.0	8,720,000	100.00

Source: U.S. Census Bureau, American Housing Survey, 2021; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Appendix C

Table C-1. AHS and Administrative Data for 2015 for Owner-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	2,754,000	68.1	414,000	10.2	3,167,000	78.3
	Septic	143,000	3.5	736,000	18.2	879,000	21.7
	Total	2,896,000	71.6	1,150,000	28.4	4,046,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2015; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table C-2. AHS and Administrative Data for 2017 for Owner-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	3,024,000	68.8	547,000	12.4	3,571,000	81.2
	Septic	75,000	1.7	750,000	17.1	825,000	18.8
	Total	3,099,000	70.5	1,297,000	29.5	4,396,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2017; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table C-3. AHS and Administrative Data for 2019 for Owner-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	3,220,000	70.4	593,000	13.0	3,813,000	83.3
	Septic	56,000	1.2	708,000	15.5	764,000	16.7
	Total	3,276,000	71.6	1,301,000	28.4	4,577,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2019; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table C-4. AHS and Administrative Data for 2021 for Owner-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	3,506,000	70.7	708,000	14.3	4,214,000	84.9
	Septic	64,000	1.3	683,000	13.8	747,000	15.1
	Total	3,569,000	71.9	1,392,000	28.1	4,961,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2021; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Appendix D

Table D-1. AHS and Administrative Data for 2015 for Renter-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,775,000	75.0	309,000	13.1	2,084,000	88.1
	Septic	75,000	3.2	207,000	8.8	282,000	11.9
	Total	1,850,000	78.2	516,000	21.8	2,366,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2015; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table D-2. AHS and Administrative Data for 2017 for Renter-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,916,000	79.5	324,000	13.4	2,240,000	92.9
	Septic	16,000	0.6	155,000	6.4	171,000	7.1
	Total	1,932,000	80.1	480,000	19.9	2,411,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2017; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table D-3. AHS and Administrative Data for 2019 for Renter-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,910,000	79.5	358,000	14.9	2,268,000	94.4
	Septic	12,000	0.5	121,000	5.0	134,000	5.6
	Total	1,922,000	80.1	479,000	19.9	2,401,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2019; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table D-4. AHS and Administrative Data for 2021 for Renter-Occupied Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,991,000	81.9	344,000	14.2	2,336,000	96.0
	Septic	11,000	0.5	85,000	3.5	96,000	4.0
	Total	2,003,000	82.4	429,000	17.6	2,432,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2021; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Appendix E

Table E-1. AHS and Administrative Data for 2015 for Vacant/URE Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,188,000	74.8	221,000	13.9	1,409,000	88.7
	Septic	16,000	1.0	164,000	10.3	179,000	11.3
	Total	1,204,000	75.8	384,000	24.2	1,588,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2015; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table E-2. AHS and Administrative Data for 2017 for Vacant/URE Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,117,000	77.6	216,000	15.0	1,334,000	92.7
	Septic	11,000	0.7	95,000	6.6	105,000	7.3
	Total	1,128,000	78.4	311,000	21.6	1,439,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2017; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.

Table E-3. AHS and Administrative Data for 2019 for Vacant/URE Housing Units in Florida.

		Administrative Data					
AHS Data		**Sewer**	**%**	**Septic**	**%**	**Total**	**Total %***
	Sewer	1,238,000	80.2	203,000	13.2	1,441,000	93.4
	Septic	9,000	0.6	93,000	6.0	101,000	6.6
	Total	1,246,000	80.8	296,000	19.2	1,542,000	100.0

Source: U.S. Census Bureau, American Housing Survey, 2019; Florida Water Management Inventory administrative records.
*Note: Details may not sum to totals due to rounding.