A Primer on Revised Estimates for the 2012 Rental Housing Finance Survey (RHFS)
February 24, 2014

Shawn Bucholtz[1]

Overview

The U.S. Department of Housing and Urban Development (HUD) and the U.S. Census Bureau have revised the 2012 Rental Housing Finance Survey (RHFS) weighting method in response to what was clearly an incorrect initial weighting method. The initial 2012 RHFS weights produced multifamily rental unit counts that were more than twice as large as the number of multifamily rental units estimated by other data sources, including the American Housing Survey (AHS), the American Community Survey (ACS), and the 2010 decennial census. This primer briefly describes the design of the sample frame, the initial weighting method, the revised weighting method, and the new estimates.

2012 RHFS Sample Frame

The RHFS is a survey of residential rental properties, where the term *property* is defined as a building or buildings that are covered under a single mortgage or owned by a single entity. The ideal RHFS sample frame is a list that includes all possible residential rental properties in the United States. Unfortunately, no such list exists, even for individual states. In fact, no known list includes all individual rental buildings.

The RHFS national sample frame was built using a list of residential addresses derived from the Census Bureau's Master Address File (MAF). A full residential address is composed of a basic street address (BSA) and, if necessary, a unit or apartment identifier. For instance, in the address "123 Main Street, Apartment 102," the "123 Main Street" portion is the BSA.

Using the MAF, in conjunction with tenure information (that is, renter or owner) from the 2010 decennial census, the Census Bureau built a list of all BSAs that were predominantly rental. The number of full addresses associated with each BSA served as a proxy for identifying the size of the rental property. It was assumed that a BSA and a property would correspond closely.

2012 RHFS Initial Weighting Issue #1: Multi-BSA Properties

Although they assumed that a BSA and a property would correspond closely, HUD and the Census Bureau anticipated finding numerous properties with multiple BSAs. Because the sample design was based on BSAs, properties with more than one BSA would have a probability of selection equal to the sum of the probabilities of selection for each BSA found on the property.

[1] Correspondence to the author can be sent to shawn.j.bucholtz@hud.gov.

The original calculation of the probability of selection for properties with multiple BSAs assumed that each BSA in the property belonged to the same stratum.[2] In fact, about 12 percent of RHFS sample properties included BSAs from different strata, which was a problem because the 2012 RHFS included eight strata, and there were large differences in the probabilities of selection from each stratum (Table 1).

Table 1. 2012 RHFS Sample Size by Stratum

Stratum	Stratum Universe Size (number of BSAs)	Stratum Sample Size (number of BSAs)	Second Stage Probability of Selection
2- to 4-unit BSAs, new construction	56,772	12	1 in 4,731
5- to 24-unit BSAs, new construction	17,685	15	1 in 1,179
25- to 49-unit BSAs, new construction	1,584	44	1 in 136
50 or more-unit BSAs, new construction	1,551	11	1 in 141
2- to 4-unit BSAs, existing	1,805,175	565	1 in 3,195
5- to 24-unit BSAs, existing	703,976	728	1 in 967
25- to 49-unit BSAs, existing	57,645	2,135	1 in 27
50 or more-unit BSAs, existing	60,320	520	1 in 116
Total		**4,030**	

BSA = basic street address. RHFS = Rental Housing Finance Survey.

To get a better sense of why problems arose when properties included BSAs from different strata, consider the following simplified example.[3] In the initial RHFS sample, a 4-unit BSA was selected from the 2- to 4-unit stratum, with a probability of selection of 1/3,195 (weight of 3,195). The interview, however, revealed an additional 48-unit BSA and a 60-unit BSA on the property. The original calculation of the probability of selection was (1/3,195 + 1/3,195 + 1/3,195 = 1/1,065), for a final property weight of 1,065. With a final property weight of 1,065, this property would represent 119,280 housing units (112 units x 1,065).

In this example, the calculation of the probability of selection for this property was clearly wrong. Recall that both of the additional BSAs had a probability of selection in our sample frame, but they were certainly not from the same stratum as the initial 4-unit BSA. The correct probability of selection for this property should have reflected the probability of selection from each of the three respective strata (2 to 4 units, 25 to 49 units, and 50 or more units). As such, the correct probability of selection is (1/3,195 + 1/27 + 1/116) = 0.04597. The revised final weight is then 21.75 (1/0.04597), and this property represents 2,436 units (112 units x 21.75), which is 116,844 fewer units than the original estimate.

[2] For sampling purposes, BSAs were stratified according to the number of MAF-unit addresses at the BSA. The BSAs in the different strata have different probabilities of selection.

[3] This example is simplified because the final RHFS weight includes a first-stage probability of selection and a nonresponse adjustment.

2012 RHFS Initial Weighting Issue #2: Units in Addresses Not Found in MAF

Another issue encountered with the RHFS was the presence of units on the property for which we had poor correspondence with BSA information in the MAF and, hence, could not accurately compute the probability of selection for the property. In some cases, not all the BSAs were collected or the BSAs collected could not be matched to the MAF. In other cases, the number of units reported by the respondent was not consistent with the number of MAF unit addresses associated with the BSAs. The net result was a probability of selection that was not consistent with the scope of the property, leading to incorrect property weights.

To get a better sense of why incorrect property weights arose, consider the following example. In the initial RHFS sample, a 4-unit BSA was selected from the 2- to 4-unit stratum, with a probability of selection of 1/3,195 (weight of 3,195). The interview revealed two additional 4-unit BSAs which had a match to a BSAS in the MAF and a 4-unit building that could not be matched to a BSA in the MAF. Still, the total unit count for the property was 16, as reported by the property owner.

The original calculation of the probability of selection was based only on the BSAs that could be matched to the MAF. In this example, three 4-unit BSAs on the property could be matched in the MAF, so the probability of selection was (1/3,195 + 1/3,195 + 1/3,195 = 1/1,065), for a final property weight of 1,065. Because the interview revealed 16 units, however, this property would represent 17,040 housing units (16 units x 1,065).

In this example, calculation of the probability of selection for this property was wrong because it did not account for the 4-unit building lacking a corresponding BSA match in the MAF. In reality it is quite likely this 4-unit building was in the MAF, but the address collection process was inaccurate or incomplete, so no match in the MAF could be found.

We chose to fix this issue by applying a "MAF unit adjustment factor" to the probability of selection. This factor is simply the count of units reported by the respondent divided by the count of MAF addresses within the verified BSAs. In our example, we were able to match three 4-unit BSAs to the MAF, whereas the respondent reported 16 units. As such, the factor would be 16/12 and the probability of selection about (1/1,065) * (16/12) = 799, meaning the property would represent 12,784 units, which is 4,256 fewer units than the original estimate.

2012 Revised Results

The two weighting mistakes described previously caused significant overcounting of housing units. The original RHFS weights produced a multifamily rental unit estimate of more than 57

million units. After making the corrections, the revised RHFS weight produced an occupied multifamily rental unit estimate of 18.47 million units.[4]

The first "smell test" for the revised estimate is to compare it with a corresponding estimate from the 2011 AHS. The AHS data allow for the exclusion of condominiums, public housing, and owner-occupied small buildings, making it an ideal data source to compare with the RHFS.[5] The 2011 AHS estimate of units in multifamily buildings (excluding condominiums, public housing, and owner-occupied small buildings) was 17.37 million.

Another "smell test" for the 2012 RHFS is to compare the number of occupied rental units in multifamily buildings with the 2011 AHS, by building size. Table 2 provides that comparison.[6] The 2012 RHFS and 2011 AHS total unit estimates are in good agreement for buildings of 2 to 9 units and for buildings with 25 to 49 units. The biggest differences appear in the 10- to 24-unit buildings and the buildings with 50 or more units.

Table 2. Comparison of 2012 RHFS and 2011 AHS Occupied Multifamily Rental Unit Counts

	2012 RHFS	2011 AHS	Percent Difference
2- to 4-unit buildings	4,472,874	4,738,252	– 5.9
5- to 9-unit buildings	3,671,189	3,841,524	– 4.6
10- to 24-unit buildings	5,821,081	4,969,144	14.6
25- to 49-unit buildings	1,582,122	1,490,999	5.8
50- to 99-unit buildings	1,353,660	1,015,400	25.0
100 or more-unit buildings	1,568,571	1,312,375	16.3
Total	**18,469,498**	**17,367,694**	**6.0**

AHS = American Housing Survey. RHFS = Rental Housing Finance Survey.

Although the 2012 RHFS and 2011 AHS estimates are not in perfect agreement, note that the 2011 AHS estimate of rental units in multifamily buildings was less than the 2011 ACS estimate by nearly 1.4 million units.[7] The 2011 AHS estimate of rental units in *any* structure type was 2 million less than the 2010 decennial census estimate. As such, good reason exists to believe the 2011 AHS underestimated the number of occupied rental units in the United States.

[4] The RHFS total unit estimate is 21,249,337. The RHFS occupied unit estimate is 18,931,537. The occupied unit estimate includes about 462,000 units in single-family structures in RHFS properties. After removing those units, the RHFS estimate of occupied units in multifamily structures is 18.47 million units.

[5] The RHFS sample excludes multifamily buildings that are condominiums or public housing, multifamily buildings of 2 to 4 units where the owner occupies one of the units, and single-family rentals.

[6] The SAS code for replicating this table is provided on the RHFS website. Users should note that the 2012 RHFS instrument only collected units by building for the first 19 building. All other units were placed into the 20th building. As such, when the respondent reported units in the 20th building, a model was built to distribute units in the 20th building into "pseudo-buildings" of various sizes. The sizes of the pseudo-buildings were based on the sizes of buildings 1 through 19. The model contains a stochastic component, so the results of the model may slightly differ from the above table. The results of table 2 are averages over 50 successive runs of the model.

[7] This comparison is of rental units in multifamily structures with two or more units. The ACS does not collect data that would permit a direct comparison with the RHFS, including whether the owner lives in a unit on the property (for 2- to 4-unit buildings), whether the building is public housing, or whether the rental unit is a condominium.

Based on the comparison with the 2011 AHS and the knowledge that the 2011 AHS rental unit estimate is less than other reputable sources, it is reasonable to conclude that the 2012 RHFS estimates based on the new weight are well within the accepted range of rental units in multifamily buildings in the United States.