American Housing Survey

Weighting Strategy for 2011–2013 CINCH Analysis

Prepared For:

U.S. Department of Housing & Urban Development
Office of Policy Development & Research

Prepared By:

Frederick J. Eggers
Fouad Moumen
Econometrica, Inc.
Bethesda, MD

Contract No. DU204SB-14-C-01
Order No. RCS-R-15-00107
Project No. 2450-008

January 2016

Table of Contents

The CINCH Objective 1

Weighting 2

New Issues With the 2011–2013 CINCH 3

SAMEDU2 4

Changes to Weighting Algorithms 5

Forward-Looking Weighting Algorithm: From 2011 to 2013 6

Backward-Looking Weighting Algorithm: From 2013 to 2011 20

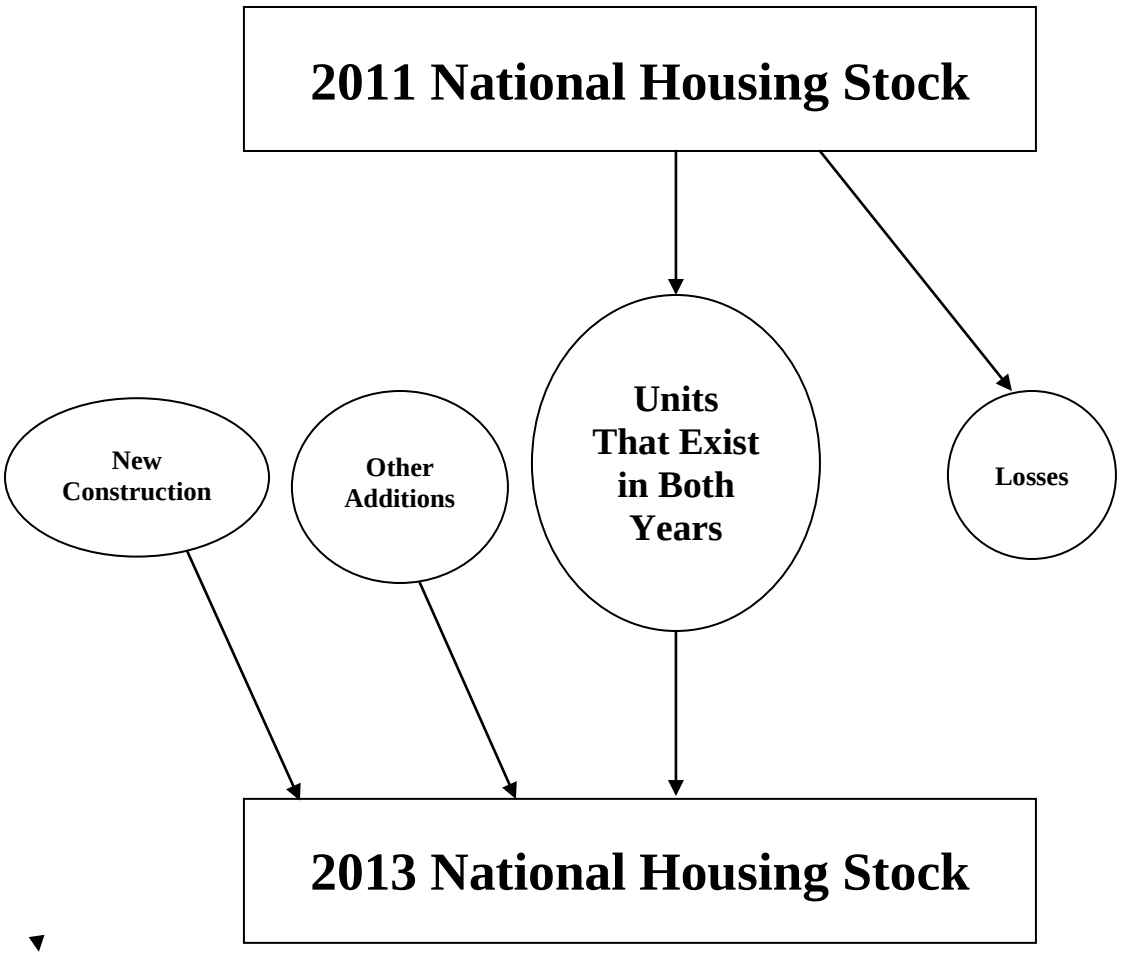
WEIGHTING STRATEGY FOR 2011–2013 CINCH ANALYSIS

This paper adapts the weighting strategy used by Econometrica, Inc., in its components of inventory change (CINCH) analysis of changes in the national housing stock between 2007 and 2009.[1]

The CINCH Objective

Figure 1 illustrates the question that CINCH analysis seeks to answer.

Figure 1: The CINCH Objective



CINCH tries to explain how the housing stock evolves from one period to the next. Figure 1 contains four ovals and two rectangles. The Census Bureau provides estimates for both rectangles and one oval (units added through new construction between 2011 and 2013). No one estimates the other three ovals: the number of units that belong to both the 2011 and 2013

[1] *Weighting Strategy for 2007–2009 CINCH Analysis* at: http://www.huduser.org/portal/datasets/cinch/cinch09/Strategy_07-09_CINCH.pdf.

housing stock, units lost to the housing stock between 2011 and 2013, and other additions to the housing stock between 2011 and 2013.

Losses can be either permanent or temporary. Units destroyed by natural disasters or intentionally demolished are permanent losses. Temporary losses include units that are condemned pending extensive repairs or units that are used for nonresidential purposes.[2] Besides new construction, additions can include units resulting from splitting up larger units, mobile home move-ins, and units that had been used formerly for nonresidential purposes.

In addition to determining the size of each oval, housing analysts find information about the characteristics of the units in the different ovals useful. Interesting characteristics include structure type, age of the unit, size of the unit, location by region, location by metropolitan status, tenure, household size and composition, resident income, and resident race and ethnicity.

CINCH analysis has three goals:

- To provide estimates for all six components of Figure 1.
- To disaggregate losses and other additions into relevant component parts.
- To characterize the units that survive from one period to the next and the units that are added or lost between periods.

The American Housing Survey (AHS) has four features that make CINCH analysis possible:

- Each unit has weights that can be used to estimate its share of the overall stock.
- The AHS tracks new construction and the various types of losses and other additions.
- The AHS has detailed information about the characteristics of each unit and its occupants.
- The AHS tracks the same unit from one period to the next so that changes in status and characteristics can be observed directly.

Weighting

Ideally, analysts would like to solve two simultaneous equations using CINCH analysis:[3]

(1) 2011 housing stock = units that exist in both years + losses.

(2) New construction + other additions + units that exist in both years = 2013 housing stock.

[2] "Potentially reversible" might be a better term than "temporary" for these types of losses.
[3] The equations are "simultaneous" because the term "units that exist in both years" appears in each equation.

Unfortunately, previous experience with CINCH analysis has shown that it is difficult to find satisfactory simultaneous solutions to the equations. For this reason, Econometrica chose to solve the two equations separately in previous CINCH studies.

Solving equation (1) is termed forward-looking analysis because it tracks what happens to the units in the 2011 housing stock. In terms of Figure 1, forward-looking analysis deals with the top rectangle and the two ovals on the right. Solving equation (2) is termed backward-looking analysis because it tracks where units in the 2013 housing stock came from. In terms of Figure 1, backward-looking analysis deals with the bottom rectangle and the three ovals on the left. In analytical terms, backward-looking analysis reverses the arrows at the bottom of Figure 1 by taking the 2013 housing stock as its starting point.

Separating the analysis into forward-looking and backward-looking components results in each observation having two weights: a weight for the forward-looking analysis (FLCINCHWT) and a weight for the backward-looking analysis (BLCINCHWT).

Solving the equations separately also results in two independent estimates of "units that exist in both years," one based on each set of weights. This paper develops algorithms to carry out the forward-looking and backward-looking analyses.

New Issues With the 2011–2013 CINCH

Oversample of HUD-assisted units

In 2011, HUD and the Census Bureau took two steps to make the AHS more useful for studying assisted rental housing. For the first time, the 2011 Public Use File (PUF) contains a variable, HUDADMIN, that identifies public housing units, units in HUD-assisted privately owned rental properties, and units whose households have HUD housing vouchers. In addition, the Census Bureau added a sample of public housing units and units in HUD-assisted privately owned rental properties to the regular AHS sample to facilitate analysis of this important subset of the housing stock by increasing the number of units available for study. The 2013 AHS included the oversample of HUD-assisted units.

The addition of the oversample complicated the construction of weights for the AHS, specifically WGT90GEO. We do not yet have reliable information on how the weights were adjusted in 2011 for the oversample, and it appears that different adjustments may have been made in the 2013 AHS. We do not know whether these adjustments create any concerns for the CINCH weighting strategy.[4]

[4] Aware that the AHS weighting process for HUD-assisted units might be different in 2011 and 2013, we eliminated a step in both the forward-looking and backward-looking algorithms.

Changes in how values for REUAD are assigned

One reason the AHS has been so valuable for CINCH analysis is that the Census Bureau tracks new construction and the various types of losses and other additions. The variable REUAD[5] indicates how new units joined the sample.

For the 2013 PUF, the Census Bureau changed the coding of REUAD. In earlier years, REUAD was based on information from the field representatives for all cases new to the sample. In 2013, the following changes were made:

1. The value of 3 (new construction) was given for cases that were added as part of a permit sample or built since the last survey year in a non-permit-issuing area.

2. The value of 4 (mobile home moved in) was set for mobile homes that were found as additional or extra units in 2013. Mobile homes in new construction were also moved to this category. Prior to the 2009–2011 CINCH analysis, new mobile homes were not classified as new construction. In 2013, the coding was changed to include new mobile homes as new construction. In the 2011–2013 CINCH, they will once again be counted as other additions.

3. There is no longer a value of 5 (house moved in). This was never a large group.

4. A value of 6 (building relisted due to structural changes) was given to additional and extra units found during the interview process that the field representatives determined to be part of this category resulting from a follow-up questionnaire about additional and extra cases. This collapses the old categories 6 (buildings relisted due to structural changes), 7 (unit created when original living quarters split into more units), and 8 (unit created when original quarters merged to fewer units).

5. Values 9 (conversion of nonresidential unit) and 10 (other, specify) were dropped. Presumably these cases are now included under the value 6.

6. A value of 11 (sample adjustment) was set for all cases that were added to the sample due to some form of sample expansion or coverage improvement.

These changes require a change in the structure of the backward-looking CINCH tables.

SAMEDU2

The AHS contains a variable to identify cases where the unit interviewed in one survey may not be the same unit that was interviewed in the previous survey. The variable (SAMEDU) takes only yes/no values. For the purpose of CINCH analysis, we created a modified version (SAMEDU2) that uses information from multiple AHS surveys to attempt to specify how the unit might differ from the unit in the previous survey. The construction of SAMEDU2 is

[5] REUAD is the "reason unit added" variable used since 1997 to track other additions to the housing stock.

explained in a companion paper.[6] In creating weights, SAMEDU2 is used to eliminate cases that may not be valid for CINCH analysis and to distinguish types of losses and additions.

In the weighting discussion, interpret SAMEDU2 as follows:

IN13_SAMEDU2 = B	Not applicable.
IN13_SAMEDU2 = 1	Not clear why SAMEDU = '2' (no).
IN13_SAMEDU2 = 2	Possibly the wrong unit was interviewed in 2011.
IN13_SAMEDU2 = 3	A new type C non-interview (a permanent loss).
IN13_SAMEDU2 = 4	Vacant mobile home lot that was occupied in 2011.
IN13_SAMEDU2 = 5	Mobile home move-in (to a vacant lot, replacing an old mobile home, or replacing a non-mobile-home structure)—note that this implies either a mobile home move-out or a demolition of another structure type.
IN13_SAMEDU2 = 6	Possible merger.
IN13_SAMEDU2 = 7	Possible split.
IN13_SAMEDU2 = 8	Possible merger or split—we cannot tell because the work has not been completed or the unit was not interviewed.

Changes to Weighting Algorithms

For the 2011–2013 CINCH, we adopted a more aggressive weighting strategy based on successful experimentation in developing weights for the backward-looking metropolitan 2009–2011 CINCH.[7]

In previous CINCH analyses, we compared estimates of various subsets of the housing stock based on CINCH weights with published estimates using AHS weights. While the two sets of estimates were generally close, we reported some deviations in estimates of race of householder, Hispanic origin of householder, metropolitan–nonmetropolitan distribution of the housing stock, and regional distribution. For the 2011–2013 CINCH, we conducted four preliminary adjustments to the weights to match published totals in these areas in the hope that the preliminary adjustments would improve the final match between the published estimates and the CINCH estimates after the last adjustments to the weights.

[6] *Listing of Programs and Variables Used in CINCH and Rental Dynamics Analysis for 2011 and 2013 American Housing Surveys.*

[7] See *Weighting Strategy for 2011 Metropolitan CINCH Analysis* by Frederick J. Eggers and Fouad Moumen at http://www.huduser.gov/portal/datasets/cinch/cinch11/Metro_Weighting_Strategy_revised.pdf.

Finally, we expanded a step suggested by a former statistician at the Census Bureau to improve estimates of mobile homes.[8] For this CINCH analysis, we control the weights in the final adjustment to equal published totals by both occupancy status (owner-occupied, renter-occupied, or vacant) and seasonal use and by structure type (single-family detached, single-family attached, structures with 2–4 units, structures with 5–19 units, structures with 20–49 units, structures with 50 or more units, and mobile homes).

Forward-Looking Weighting Algorithm: From 2011 to 2013

The following are the steps necessary to prepare the data to analyze what happened between 2011 and 2013 to units that existed in 2011. AHS variables are given their codebook names and presented in capital letters. We refer to 2011 variables by the prefix IN11_; 2013 variables are labeled IN13_.

1. Use the 2009, 2011, and 2013 PUFs to create SAMEDU2 for any units in the 2013 PUF that have SAMEDU = '2'—see *Listing of Programs and Variables Used in CINCH and Rental Dynamics Analysis for 2011 and 2013 American Housing Surveys* for the construction of SAMEDU2.

   {Dav, this draft includes the revised coding for SAMEDU2 because the referenced paper is still being written. – Fred}

   IN13_SAMEDU2 = B

   IF IN13_SAMEDU = '2', IN13_SAMEDU2 = 1

   IF (IN13_SAMEDU = '2' AND (((IN09_NUNIT2 = IN13_NUNIT2) AND (IN11_NUNIT2 NE IN13_NUNIT2)) OR ((IN09_ROOMS = IN13_ROOMS) AND (IN11_ROOMS NE IN13_ROOMS)))) THEN IN13_SAMEDU2 = 2

   IF (IN13_SAMEDU = '2' AND ((IN11_NOINT = 'B' OR IN11_NOINT LT 30) AND IN13_NOINT GE 30)) THEN IN13_SAMEDU2 = 3

   IF (IN13_SAMEDU = '2' AND (IN11_NUNIT2 = 4 AND IN13_NOINT = 13)) THEN IN13_SAMEDU2 = 4

   IF (IN13_SAMEDU = '2' AND IN13_SAMEDU2 NE 2 AND (IN13_NUNIT2 = 4 AND (IN11_NUNIT2 = B OR (IN11_NUNIT2 = 4 AND (IN11_BUILT NE IN13_BUILT)) OR IN11_NUNIT2 LT 4))) THEN IN13_SAMEDU = 5

   IF (IN13_SAMEDU = '2' AND IN13_NUNIT2 NE 4 AND (IN13_BUILT LT 2011 AND ((IN09_ROOMS = IN11_ROOMS) AND (IN11_ROOMS LT IN13_ROOMS))) AND IN13_NOINT = B) THEN IN13_SAMEDU2 = 6

[8] See page 6 of *Weighting Strategy for 2003–2005 CINCH Analysis* by Frederick J. Eggers at http://www.huduser.gov/portal/datasets/cinch/cinch05/CINCHWeightingStrategy2007.pdf.

IF (IN13_SAMEDU = '2' AND IN13_NUNIT2 NE 4 AND (IN13_BUILT LT 2011 AND ((IN09_ROOMS = IN11_ROOMS) AND (IN11_ROOMS GT IN13_ROOMS))) AND IN13_NOINT = B) THEN IN13_SAMEDU2 = 7

IF (IN13_SAMEDU = '2' AND IN13_NUNIT2 NE 4 AND (IN13_BUILT LT 2011 AND ((IN09_ROOMS = IN11_ROOMS) AND (IN11_ROOMS NE IN13_ROOMS))) AND 1 LE IN13_NOINT LE 12) THEN IN13_SAMEDU2 = 8

IN13_SAMEDU	**IN13_SAMEDU2**								
	1	**2**	**3**	**4**	**5**	**7**	**8**	**B**	**Total**
1	0	0	0	0	0	0	0	67,659	67,659
2	31	7	280	32	10	1	2	0	363
Total	31	7	280	32	10	1	2	67,659	68,022

2. Merge the 2011 and 2013 files, using the flat files.

    a. Eliminate non-matches.

A: IN BOTH 11 & 13	68,022
B: IN 11 ONLY	118,426
C: IN 13 ONLY	16,333

    b. Test to see if there are any cases in the matched sample where IN11_NATLFLAG = '2' (part of the metropolitan sample in 2011). If there are such units, we may have to adjust the pure weight (PWT) for these units. No cases with IN1_NATLFLAG = '2' were found.

    c. Test to see if there are any cases in the matched sample that are part of the special oversample of HUD-assisted units (IN11_HUDSAMP = '1'), which began in 2011. If there are such units, we may have to adjust PWT for these units. Count the number of these cases. There are 4,208 cases with IN11_HUDSAMP = '1'

d. Do an unweighted frequency distribution of IN13_NOINT.

IN13_NOINT	Frequency	Percent	Cumulative frequency	Cumulative percent
B	56,780	83.47	56,780	83.47
1	952	1.40	57,732	84.87
2	89	0.13	57,821	85.00
3	6,977	10.26	64,798	95.26
4	41	0.06	64,839	95.32
5	68	0.10	64,907	95.42
6	1,399	2.06	66,306	97.48
10	10	0.01	66,316	97.49
11	26	0.04	66,342	97.53
12	219	0.32	66,561	97.85
13	165	0.24	66,726	98.09
14	275	0.40	67,001	98.50
15	64	0.09	67,065	98.59
16	166	0.24	67,231	98.84
17	87	0.13	67,318	98.97
30	272	0.40	67,590	99.36
31	88	0.13	67,678	99.49
32	39	0.06	67,717	99.55
33	35	0.05	67,752	99.60
36	5	0.01	67,757	99.61
37	250	0.37	68,007	99.98
40	15	0.02	68,022	100.00

e. Eliminate cases where IN13_NOINT GE 38. This eliminates losses due to sample changes. CINCH should ignore these losses because they are not physical losses and because we cannot say anything useful about what happens to them. (15 cases)

f. Eliminate cases where 1 LE IN13_SAMEDU2 LE 2. This eliminates cases where it is possible that the Census Bureau went to the wrong unit in 2011. (38 cases)

3. Do an unweighted frequency distribution of IN11_NOINT.

IN11_NOINT	Frequency	Percent	Cumulative frequency	Cumulative percent
B	58,367	85.87	58,367	85.87
1	705	1.04	59,072	86.91
2	43	0.06	59,115	86.97
3	6,549	9.64	65,664	96.61
4	68	0.10	65,732	96.71
5	34	0.05	65,766	96.76
6	1,006	1.48	66,772	98.24
10	23	0.03	66,795	98.27
11	64	0.09	66,859	98.37
12	287	0.42	67,146	98.79
13	181	0.27	67,327	99.06
14	290	0.43	67,617	99.48
15	62	0.09	67,679	99.57
16	196	0.29	67,875	99.86
17	94	0.14	67,969	100.00

Eliminate all observations that were 2011 type B or type C losses (10 LE IN11_NOINT). These units were not part of the 2011 stock and therefore are not tracked in the forward-looking analysis. Note that because of the changed treatment of type C losses in PUFs beginning with the 2011 survey, merging and keeping only matches eliminates any type C units from 2011. (1,197 cases)

4. Adjust PWTs for 2011 in 28 metropolitan areas surveyed as part of the metropolitan AHS.

   In 2011, the AHS combined the national and metropolitan surveys. Twenty-eight metropolitan areas have sample cases from the national sample and the metropolitan sample. The cases from the metropolitan sample cannot be used in the national CINCH as they have no 2013 matches. For the 28 areas, each case has 4 weights in 2011: PWT, an adjusted weight to be used in the national analysis (WGT90GEO), an adjusted weight to be used for the metropolitan analysis (WGTMETRO), and an adjusted weight to be used for national analysis *if* only national cases are used (PUFWGT). CINCH weights are based on PWTs. For these areas, PWT takes into account both samples in 2011 and therefore is smaller than what we would desire it to be.

   The following table uses only cases from the national sample. It reports the average ratio of (PWT in 2009)/(PWT in 2011). The table demonstrates that the PWTs for 2011 for cases in the 28 metropolitan areas are lower than the typical national case, and the ratio is very consistent for each area.

Area	Sample size	Ratio of (PWT in 2009)/(PWT in 2011)			
		Mean	90th per	75th per	Median
Anaheim	378	2.242302	2.15324	2.15324	2.15324
Atlanta	359	2.262804	2.16734	2.16734	2.16734
Birmingham	98	2.645409	2.27609	2.27609	2.27609
Buffalo	161	2.197836	2.19905	2.19905	2.19905
Cincinnati	195	2.986955	2.39809	2.39809	2.39809
Cleveland	282	2.441410	2.20399	2.20399	2.20399
Columbus	218	2.171913	2.17280	2.17280	2.17280
Dallas	445	2.195513	2.12841	2.12841	2.12841
Denver	124	2.457298	1.97604	1.97604	1.97604
Fort Worth	234	2.154470	2.15694	2.15694	2.15694
Indianapolis	180	3.188909	2.20961	2.20961	2.20961
Kansas City	244	2.346076	1.98021	1.98021	1.98021
Los Angeles	1,335	3.376580	3.04836	3.04836	3.04836
Memphis	144	2.616419	2.38010	2.38010	2.38010
Milwaukee	211	2.376914	2.20947	2.20947	2.20947
New Orleans	173	2.344619	2.34825	2.34825	2.34825
Oakland	356	2.601272	2.05593	2.05593	2.05593
Phoenix	506	2.196403	2.13824	2.13824	2.13824
Pittsburgh	302	2.498695	2.15729	2.15729	2.15729
Portland	262	2.148079	1.91498	1.91498	1.91498
Providence	196	2.885895	2.05373	2.05373	2.05373
Riverside	266	2.272109	2.05198	2.05198	2.05198
Sacramento	222	2.647437	2.17844	2.17844	2.17844
San Diego	417	2.361211	2.19389	2.19389	2.19389
San Francisco	297	2.332360	2.10562	2.10562	2.10562
San Jose	248	2.285886	2.14742	2.14742	2.14742
St. Louis	323	3.025786	2.17598	2.17598	2.17598
Virginia Beach	250	2.879743	2.19378	2.19378	2.19378
Rest of sample	46,639	1.258179	1.15233	1.15233	1.15233

a. Adjust PWT in each of the 28 metropolitan areas as follows:

   If case is not in one of 28 areas: IN11_ADJPWT = IN11_PWT

   If case is in one of 28 areas: IN11_ADJPWT = (median from above table)*IN11_PWT.

b. MXPWT = IN11_ADJPWT (Note: We dropped the old step 5 where MXPWT = max (IN13_PWT, IN11_ADJPWT) because of the change in PWT between surveys for HUD-assisted units.)

5. Obtain from the Census Bureau tables an estimate of the 2011 stock (BASECOUNT = 132,419,000).

6. Compute SMXPWT = sum of MXPWT after step 5; this sum is a first estimate of the size of the housing stock based on the units retained for analysis. SMXPWT = 125,303,787, based on 66,772 cases.

7. Compute FLCINCHWT = MXPWT*(BASECOUNT/SMXPWT). This computation ratios the weights up so that they sum to the 2011 stock. BASECOUNT/SMXPWT = 1.0567837

8. Identify *sames, losses,* and *interviewed losses:*

    a. SAME = 1 if IN11_ISTATUS = 1, 2, or 3 AND IN13_ISTATUS = 1, 2, or 3 AND NOT(IN13_SAMEDU2 GE 4) (57,277 cases)

    b. LOSS = 1 if IN11_ISTATUS = 1, 2, 3, or 4 AND (10 LE IN13_NOINT LT 38 OR IN13_SAMEDU2 GE 4). IN13_SAMEDU2 GE 4 means that the Census Bureau considers this a different unit than the unit in the 2011 sample and, therefore, we will treat the 2011 unit as a loss. (845 cases)

    c. INTLOSS = 1 if IN11_ISTATUS = 1, 2, or 3 AND LOSS = 1 (749 cases)

9. Calculate:

    a. SSAME = sum of FLCINCHWT for all SAME = 1 SSAME = 103,066,125

    b. SLOSS = sum of FLCINCHWT for all LOSS = 1 SLOSS = 1,710,733

    c. SINTLOSS = sum of FLCINCHWT for INTLOSS = 1 SINTLOSS = 1,593,445

10. For CINCH analysis, we need information on the characteristics of units and their occupants in both 2011 and 2013 for all units that were part of the stock in both 2011 and 2013. For units that are part of the stock in only 2011, we need information on the characteristics of the units and their occupants only in 2011. Up to this point, we retained units that failed to meet these conditions so that we can get good estimates of the number of losses (SLOSS).

    Keep for future analysis only those units where SAME = 1 OR INTLOSS = 1.

    Note that this formulation keeps a few 2013 type A non-interviews if the unit is interviewed in 2011 *and* is also an eligible SAMEDU = '2' case. Since we treat the 2013 version of the unit as a different unit, we do not need to know the characteristics of the unit or its occupants in 2013 for the forward-looking analysis.

11. Calculate:

    a. Ratio1 = (BASECOUNT – SLOSS)/SSAME = 1.2681981

    b. Ratio2 = SLOSS/SINTLOSS = 1.0736066

12. Recalculate FLCINCHWT as follows:

    a. For SAME = 1, FLCINCHWT = Ratio1*FLCINCHWT

    b. For INTLOSS = 1, FLCINCHWT = Ratio2*FLCINCHWT

13. Do a preliminary adjustment to FLCINCHWT to improve counts of householders by race.

    a. From published reports, obtain estimated 2011 counts for units by race.

White alone	92,820,000
Black alone	14,694,000

    b. Develop estimates for these same categories using FLCINCHWT with these formulas:

White alone	IN11_ISTATUS = '1' AND IN11_RACE1 = '01'	90,018,331
Black alone	IN11_ISTATUS = '1' AND IN11_RACE1 = '02'	13,823,846

    c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

    For example, if the estimate in step b for units with “White only” householders is 90,018,331 units, then the ratio for the top cell in step c is 1.03112.

White alone	1.03112
Black alone	1.06295

    d. Adjust FLCINCHWT by applying the new adjustment ratios to existing FLCINCHWT using the formulas in step b to determine which FLCINCHWT to adjust by which ratio. Calculate the sum of FLCINCHWT by category.

White alone	92,820,000
Black alone	14,694,000

14. Do a second adjustment to FLCINCHWT to improve count of householders by ethnicity.

    a. From published reports, obtain estimated 2011 count for units by ethnicity.

Hispanic	13,841,000

    b. Develop estimate for this category using FLCINCHWT with this formula:

Hispanic	IN11_ISTATUS = '1' AND IN11_SPAN1 = '01'	14,983,978

c. Create new adjustment ratio by taking the ratio of the published number in step a to the estimate in step b. Ratio = 0.92372

d. Adjust FLCINCHWT by applying the new adjustment ratio to existing FLCINCHWT using the formula in step b to determine which FLCINCHWT to adjust. Calculate the sum of FLCINCHWT by category = 13,841,000.

15. Do a third adjustment to FLCINCHWT to improve regional counts.

a. From published reports, obtain estimated 2011 counts for units by region.

Northeast	23,717,000
Midwest	29,545,000
South	50,381,000
West	28,776,000

b. Develop estimates for these same categories using FLCINCHWT with these formulas:

Northeast	REGION = '1'	23,321,666
Midwest	REGION = '2'	31,253,062
South	REGION = '3'	51,294,317
West	REGION = '4'	29,078,800

c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

For example, if the estimate in step b for Northeast units is 23,321,666, then the ratio for the top cell in step c is 1.01695.

Northeast	1.01695
Midwest	0.94535
South	0.98219
West	0.98959

d. Adjust FLCINCHWT by applying the new adjustment ratios to existing FLCINCHWT using the formulas in step b to determine which FLCINCHWT to adjust by which ratio.

For example, for units in the Northeast, create a revised FLCINCHWT by applying the formula:

FLCINCHWT = 1.01695*FLCINCHWT

to all FLCINCHWT values of units where REGION = '1'

Enter the sum of FLCINCHWT into the following matrix:

Northeast	23,717,000
Midwest	29,545,000
South	50,381,000
West	28,776,000

16. Do a fourth adjustment to FLCINCHWT to improve metropolitan status counts.

    a. From published reports, obtain estimated 2011 counts for units by metropolitan status.

Central city	38,599,000
Suburb	65,418,000
Nonmetro	28,402,000

    b. Develop estimates for these same categories using FLCINCHWT with these formulas:

Central city	METRO3 = '1'	40,519,628
Suburb	METRO3 = {'2','3'}	33,283,138
Nonmetro	METRO3 = {'4','5'}	58,616,234

    c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

       For example, if the estimate in step b for units in central cities is 40,519,628 units, then the ratio for the top cell in step c is 0.95260.

Central city	0.95260
Suburb	1.11604
Nonmetro	0.85335

    d. Adjust FLCINCHWT by applying the new adjustment ratios to existing FLCINCHWT using the formulas in step b to determine which FLCINCHWT to adjust by which ratio.

       For example, for units in the central cities, create a revised FLCINCHWT by applying the formula:

       FLCINCHWT = 0.95260*FLCINCHWT

       to all FLCINCHWT values of units where METRO3 = '1'

Enter the sum of FLCINCHWT into the following matrix:

Central city	38,599,000
Suburb	65,418,000
Nonmetro	28,402,000

17. Adjust the FLCINCHWT to sum to the published totals for 2011 for units by unit type and occupancy status.

    a. From published reports, obtain estimated 2011 counts for units by unit type and occupancy status.

    The published numbers are:

	Owner-occupied	**Renter-occupied**	**Vacant**	**Seasonal**
Single-family detached	62,662,000	11,099,000	6,664,000	2,549,000
Single-family attached	4,090,000	2,654,000	798,000	226,000
2–4 unit structures	1,419,000	7,537,000	1,535,000	187,000
5–19 unit structures	1,101,000	9,341,000	1,712,000	228,000
20+ unit structures	1,142,000	6,672,000	1,439,000	316,000
Mobile homes	5,678,000	1,512,000	1,233,000	626,000

b. Develop estimates for these same categories using FLCINCHWT with these formulas:

	Owner-occupied	**Renter-occupied**	**Vacant**	**Seasonal**
Single-family detached	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '1'	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) AND IN11_NUNIT2 = '1'	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '1'	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '1'
Single-family attached	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '2'	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) AND IN11_NUNIT2 = '2'	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '2'	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '2'
2–4 unit structures	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '3' AND IN11_NUNITS = {2,3,4}	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) IN11_NUNIT2 = '3' AND IN11_NUNITS = {2,3,4}	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) IN11_NUNIT2 = '3' AND IN11_NUNITS = {2,3,4}	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) IN11_NUNIT2 = '3' AND IN11_NUNITS = {2,3,4}
5–19 unit structures	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '3' AND 5 LE IN11_NUNITS LE 19	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) AND IN11_NUNIT2 = '3' AND 5 LE IN11_NUNITS LE 19	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '3' AND 5 LE IN11_NUNITS LE 19	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '3' AND 5 LE IN11_NUNITS LE 19
20+ unit structures	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '3' AND 20 LE IN11_NUNITS	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) AND IN11_NUNIT2 = '3' AND 20 LE IN11_NUNITS	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '3' AND 20 LE IN11_NUNITS	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '3' AND 20 LE IN11_NUNITS
Mobile homes	IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '4'	IN11_ISTATUS = “1” AND (2 LE IN11_TENURE LE 3) AND IN11_NUNIT2 = '4'	IN11_ISTATUS = {'2','3'} AND NOT(8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '4'	IN11_ISTATUS = {'2','3'} AND (8 LE IN11_VACANCY LE 11) AND IN11_NUNIT2 = '4'

The estimates are:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	60,104,119	11,143,050	7,946,013	2,158,779
Single-family attached	3,980,315	2,644,268	1,009,682	246,029
2–4 unit structures	1,378,737	7,705,253	1,994,606	177,471
5–19 unit structures	1,119,958	9,727,534	2,425,585	270,822
20+ unit structures	1,213,177	7,312,820	2,007,250	390,423
Mobile homes	4,577,721	1,231,509	1,199,181	454,697

c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

   For example, if the estimate in step b for owner-occupied single-family detached units is 60,104,119, then the ratio for the upper left cell in step c is 1.04256.

   The ratios are:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	1.04256	0.99605	0.83866	1.18076
Single-family attached	1.02756	1.00368	0.79035	0.91859
2–4 unit structures	1.02920	0.97816	0.76958	1.05369
5–19 unit structures	0.98307	0.96026	0.70581	0.84188
20+ unit structures	0.94133	0.91237	0.71690	0.80938
Mobile homes	1.24036	1.22776	1.02820	1.37674

d. Adjust FLCINCHWT by applying the new adjustment ratios to existing FLCINCHWT using the formulas in step b to determine which FLCINCHWT to adjust by which ratio.

   For example, for owner-occupied single-family detached units, create a final FLCINCHWT by applying the formula:

   FLCINCHWT = 1.04256*FLCINCHWT

   to all FLCINCHWT values of units where IN11_ISTATUS = “1” AND IN11_TENURE = 1 AND IN11_NUNIT2 = '1'

Enter the sum of FLCINCHWT into the following matrix:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	62,662,000	11,099,000	6,664,000	2,549,000
Single-family attached	4,090,000	2,654,000	798,000	226,000
2–4 unit structures	1,419,000	7,537,000	1,535,000	187,000
5–19 unit structures	1,101,000	9,341,000	1,712,000	228,000
20+ unit structures	1,142,000	6,672,000	1,439,000	316,000
Mobile homes	5,678,000	1,512,000	1,233,000	626,000

At the end of the weighting process, we have created weights (FLCINCHWT) that reproduce exactly the published counts for occupancy status by structure type.

18. In steps 13, 14, 15, and 16, we successively adjusted the weights to reproduce exactly the count of householders by race, the count of householders by ethnicity, the count of units by region, and the count of units by metropolitan/nonmetropolitan status. The final adjustment reproduces exactly the published counts for occupancy status by structure type. However, each step has disturbed the match achieved by the previous steps so that by the end the only exact match is occupancy status by structure type.

    Now we compare estimates to published counts to see how closely the final weights estimate counts by race, ethnicity, region, and metropolitan/nonmetropolitan status.

Category	Formula	Estimate	Published	Percent error
Region				
Northeast	REGION = '1'	23,977,582	23,717,000	1.1
Midwest	REGION = '2'	29,209,333	29,545,000	-1.1
South	REGION = '3'	50,236,671	50,381,000	-0.3
West	REGION = '4'	28,996,413	28,776,000	0.8
Metropolitan				
Central city	METRO3 = '1'	37,400,153	38,599,000	-3.1
Suburb	METRO3 = {'2','3'}	65,872,038	65,418,000	0.7
Nonmetro	METRO3 = {'4','5'}	29,147,809	28,402,000	2.6
Race/ethnicity				
White only	RACE1 = '01'	93,136,989	92,820,000	0.3
Black only	RACE1 = '02'	14,415,727	14,694,000	-1.9
Hispanic	SPAN1 = '01'	13,973,879	13,841,000	1.0

The stepwise adjustment process has achieved close fits for these key categories.

19. Next we derive estimates of units in public housing and units in HUD-assisted, privately owned properties to see if there are any major discrepancies in this important area. We compare these estimates first to estimates based on AHS weights and then to numbers from the HUD budget.

	CINCH estimates	**AHS estimates**	**Percent error**	**HUD budget**	**Percent error**
Public housing	1,053,471	886,616	18.8	1,033,000	2.00
Vouchers	2,111,468	1,817,042	16.2	2,086,000	1.20
Privately owned	1,334,731	1,113,262	19.9	1,419,000	-5.90

The comparison between estimates based on CINCH weights and AHS weights is the legitimate test. Our weights appear to inflate the count of HUD-assisted units as reported by the AHS. Interestingly the forward-looking CINCH estimates are much closer to the HUD budget estimates.

20. As a final check, compare FLCINCHWT to IN11_PWT and IN11_PUFWGT to see if there are any obvious problems in the structure of the weights.

    a. Create two new variables, FLRATPWT and FLRATWGT, as follows:

    IF MXPWT = 0 THEN FLRATPWT = "."

    IF MXPWT GT 0 THEN FLRATPWT = FLCINCHWT/MXPWT

    IF IN11_PUFWGT = 0 THEN FLRATWGT = "."

    IF IN11_PUFWGT GT 0 THEN FLRATWGT = FLCINCHWT/ IN11_PUFWGT

b. Determine means, mode, medians, and key percentiles for the distribution of MXPWT, IN11_WGT90GEO, FLCINCHWT, FLRATPWT, and FLRATWGT.

	MXPWT	IN11_PUFWGT	FLCINCHWT	Ratio FLCINCHWT/ MXPWT	Ratio FLCINCHWT/ PUFWGT
Mean	1,866.63	2,217.12	2,495.85	1.33268	1.26892
Mode	2,046.60	1,652.62	3,232.17	1.57929	1.12288
Minimum	12.05	0.00	13.59	0.65616	1.02557
1 percentile	126.43	113.07	135.90	0.85185	0.42258
5 percentile	126.43	166.45	167.55	0.94207	0.63036
10 percentile	773.65	618.77	919.91	1.06112	0.73029
25 percentile	2,046.60	2,732.93	2,136.03	1.18856	0.92480
50 percentile	2,046.60	2,338.71	2,604.77	1.32217	1.13311
75 percentile	2,046.60	2,732.93	3,110.91	1.52004	1.38253
90 percentile	2,358.36	3,149.44	3,346.54	1.59117	1.79849
95 percentile	2,358.36	3,628.11	3,612.62	1.63517	2.29845
99 percentile	3,369.09	5,096.69	4,540.47	1.87891	3.73057
Maximum	26,264.64	23,253.44	40,778.00	2.03779	19.02400

We would expect FLCINCHWT to be larger than either MXPWT or PUFWGT because we eliminated a large number of cases from the 2011 PUF at step 10. Nothing alarming stands out from this overview of FLCINCHWT. We suspect the high maximum and 90–99 percentiles for the ratio of FLCINCHWT to PUFWGT result from the low final weights assigned to some units in the HUD oversample. We are satisfied with the ratio of FLCINCHWT to PUFWGT; its interquartile range is 0.92480 to 1.38253.

Backward-Looking Weighting Algorithm: From 2013 to 2011

The following are the steps necessary to prepare the data to analyze what happened between 2011 and 2013 to units that existed in 2013. AHS variables are given their codebook names and presented in capital letters. We refer to 2013 variables by the prefix IN13_; 2011 variables are labeled IN11_.

1. Merge the 2013 and 2011 files, using the flat files.

    a. Eliminate cases not in the 2013 PUF—that is, the merged file should consist of units found in 2013, whether or not they are in 2011.

A: IN BOTH 11 & 13	68,022
B: IN 11 ONLY	118,426
C: IN 13 ONLY	16,333
A + C	84,355

b. Test to see if there are any cases in the matched sample where IN11_NATLFLAG = '2' (part of the metropolitan sample in 2011). If there are such units, we may have to adjust PWT for these units. No cases in A have IN11_NATLFLAG = '2'

c. Test to see if there are any cases in the matched sample that are part of the special oversample of HUD-assisted units (IN11_HUDSAMP = '1'), which began in 2011. If there are such units, we may have to adjust PWT for these units. (4,208 cases with IN11_HUDSAMP = '1')

d. Do unweighted frequency distributions of IN13_REUAD and IN13_NOINT.

IN13_REUAD	Frequency	Percent
B	74,100	87.84
D	2	0.00
3	738	0.87
4	10	0.01
6	1	0.00
7	37	0.04
8	4	0.00
9	13	0.02
10	37	0.04
11	9,413	11.16

IN13_NOINT	Frequency	Percent
B	70,044	83.0
1	1,227	1.5
2	101	0.1
3	8,576	10.2
4	71	0.1
5	87	0.1
6	1,753	2.1
10	20	0.0
11	105	0.1
12	284	0.3
13	197	0.2
14	308	0.4
15	84	0.1
16	204	0.2
17	106	0.1
30	442	0.5
31	100	0.1
32	69	0.1
33	59	0.1
36	9	0.0
37	478	0.6
38	5	0.0
40	26	0.0
Total	84,355	100.0

2. Eliminate:

    a. IN13_NOINT GE 10. These are type B or type C losses in 2013. These units are not part of the 2013 stock, and therefore we do not track them backwards. (2,496 cases)

    b. 1 LE IN13_SAMEDU2 LE 2. These are cases where it is possible that the Census Bureau interviewed a different unit in 2013. (16 cases)

    c. IN13_REUAD = 11. These are cases added as sample adjustments. They are part of the 2013 housing stock, but we cannot tell whether they were in the 2011 stock or added by new construction or other means between 2011 and 2013. (8,921 cases after previous deletions—these cases were part of the metropolitan oversample for the 5 large metropolitan areas. Total cases after step 2: 72,922)

3. MXPWT = IN13_PWT (Note: We dropped the old step 3 where MXPWT = max (IN11_PWT, IN13_PWT) because of the widespread change in PWT between surveys for HUD-assisted units. This step was used in the past to adjust for a few idiosyncratic changes in PWT between surveys.)

4. Obtain from the Census Bureau tables an estimate of the 2013 stock (CURRENTCOUNT = 132,832,000).

5. Compute SMXPWT = sum of MXPWT after step 3; this sum (120,668,913) is a first estimate of the size of the 2013 housing stock based on units retained for analysis.

6. Compute BLCINCHWT = MXPWT*(CURRENTCOUNT/SMXPWT) = MXPWT * 1.1007972. This computation ratios the weights up so that they sum to the 2013 stock.

7. Identify *sames, new construction, interviewed new construction, other adds,* and *interviewed other adds.*[9]

    a. SAME = 1 if IN11_ISTATUS = 1, 2, or 3 AND IN13_ISTATUS = 1, 2, OR 3 AND NOT(IN13_SAMEDU2 GE 4) AND NOT(IN13_NUNIT2 = '4' AND IN13_BUILT GE 2011 AND IN11_BUILT NE 2011) (52,276)

    b. NC = 1 if IN13_ISTATUS = 1, 2, 3, or 4 AND ((IN13_REUAD = 3) OR (10 LE IN11_NOINT LE 11) OR (IN13_NUNIT2 = '4' AND IN13_BUILT GE 2011 AND IN11_BUILT NE 2011)) AND NOT(IN13_SAMEDU2 GE 4) (687)

    c. INTNC = 1 IF NC = 1 AND IN13_ISTATUS = 1, 2, or 3 (561)

    d. ADD = 1 if IN13_ISTATUS = 1, 2, 3, or 4 AND ((4 LE IN13_REUAD LE 10) OR (12 LE IN11_NOINT LE 17) OR (IN13_SAMEDU2 GE 4)) AND NOT NC = 1 (398)

    e. INTADD = 1 if ADD = 1 AND IN13_ISTATUS = 1, 2, OR 3 (368)

    (19,561 cases were not SAME = 1 because they were in the stock in both years but were not interviewed in both years.)

8. Calculate:

    a. SSAME = sum of BLCINCHWT for all SAME = 1 (100,834,548)

    b. SNC = sum of BLCINCHWT for NC = 1 (1,175,352)

        i. SNCMH = sum of BLCINCHWT for NC = 1 AND IN13_NUNIT2 = 4 (16,227)

[9] *Other adds* are units that were type B losses in 2011 but are in the 2013 housing stock, plus new housing units that are not new construction, such as the conversion to residential use of a warehouse or mobile home move-in.

ii. SNCOTH = sum of BLCINCHWT for NC = 1 AND IN13_NUNIT2 NE 4 (1,159,125)

c. SINTNC = sum of BLCINCHWT for INTNC = 1 (966,586)

i. SINTNCMH = sum of BLCINCHWT for INTNC = 1 AND IN13_NUNIT2 = 4 (16,227)

ii. SINTNCOTH = sum of BLCINCHWT for INTNC = 1 AND IN13_NUNIT2 NE 4 (950,359)

d. SADD = sum of BLCINCHWT for ADD = 1 (766,953)

e. SINTADD = sum of BLCINCHWT for INTADD = 1 (717,891)

9. Calculate:

a. Ratio1 = (CURRENTCOUNT – (SADD + SNC))/SSAME = 1.2980640

b. Ratio2 = SNCMH/SINTNCMH = 1

c. Ratio3 = SNCOTH/SINTNCOTH = 1.2196709

d. Ratio4 = SADD/SINTADD = 1.0683428

10. Keep units that are SAME = 1 OR INTNC = 1 OR INTADD = 1 (53,205 cases)

For CINCH analysis, we need information on the characteristics of units and their occupants in both 2011 and 2013 for all units that were part of the stock in both 2011 and 2013. For units that are part of the stock in only 2013, we need information on the characteristics of the units and their occupants only in 2013. Up to this point, we retained units that failed to meet these conditions so that we can get good estimates of the number of newly constructed units (SNC) and other additions (SADD). This step eliminated units (units that were in both the 2011 and 2013 stock and new units), leaving a sample of 53,205 units.

11. Recalculate BLCINCHWT as follows:

a. For SAME = 1, BLCINCHWT = Ratio1*BLCINCHWT

b. For INTNC = 1 AND IN13_NUNIT2 = 4, BLCINCHWT = Ratio2*BLCINCHWT

c. For INTNC = 1 AND IN13_NUNIT2 NE 4, BLCINCHWT = Ratio3*BLCINCHWT

d. For INTADD = 1, BLCINCHWT = Ratio4*BLCINCHWT

12. Do a preliminary adjustment to BLCINCHWT to improve counts of householders by race.

    a. From published reports, obtain estimated 2013 counts for units by race and ethnicity of the householder.

White alone	93,298,000
Black alone	15,023,000

    b. Develop estimates for these same categories using BLCINCHWT with these formulas:

White alone	IN13_ISTATUS = '1' AND IN13_RACE1 = '01'	91,343,530
Black alone	IN13_ISTATUS = '1' AND IN13_RACE1 = '02'	13,081,274

    c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

    For example, if the estimate in step b for units with “White only” householders is 91,343,530 units, then the ratio for the top cell in step c is 1.02140.

White alone	1.02140
Black alone	1.14844

    d. Adjust BLCINCHWT by applying the new adjustment ratios to existing BLCINCHWT using the formulas in step b to determine which BLCINCHWT to adjust by which ratio.

    For example, for units with “White only” householders, create a revised BLCINCHWT by applying the formula:

    BLCINCHWT = 1.02140*BLCINCHWT

    to all BLCINCHWT values of units where IN13_ISTATUS = '1' AND IN13_RACE1 = '01'

    Enter the sum of BLCINCHWT into the following matrix:

White alone	93,298,000
Black alone	15,023,000

13. Do a second adjustment to BLCINCHWT to improve count of householders by ethnicity.

    a. From published reports, obtain estimated 2013 count for units by ethnicity.

Hispanic	14,681,000

    b. Develop estimate for this category using BLCINCHWT with this formula:

Hispanic	IN13_ISTATUS = '1' AND IN13_SPAN1 = '1'	14,993,557

    c. Create new adjustment ratio by taking the ratio of the published number in step a to the estimate in step b.

    14,681,000/14,993,557 = 0.97915

    d. Adjust BLCINCHWT by applying the new adjustment ratio to existing BLCINCHWT using the formula in step b to determine which BLCINCHWT to adjust.

    BLCINCHWT = 0.97915*BLCINCHWT

    for all BLCINCHWT values of units where IN13_ISTATUS = '1' AND IN13_SPAN1 = '1'

Hispanic	14,681,000

14. Do a third adjustment to BLCINCHWT to improve regional counts.

    a. From published reports, obtain estimated 2013 counts for units by region.

Northeast	23,719,000
Midwest	29,606,000
South	50,679,000
West	28,828,000

    b. Develop estimates for these same categories using BLCINCHWT with these formulas:

Northeast	REGION = '1'	19,065,453
Midwest	REGION = '2'	29,586,254
South	REGION = '3'	56,563,725
West	REGION = '4'	31,200,208

c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

   For example, if the estimate in step b for Northeast units is 19,065,453, then the ratio for the top cell in step c is 1.24408.

Northeast	1.24408
Midwest	1.00067
South	0.89596
West	0.92397

d. Adjust BLCINCHWT by applying the new adjustment ratios to existing BLCINCHWT using the formulas in step b to determine which BLCINCHWT to adjust by which ratio.

   For example, for units in the Northeast, create a revised BLCINCHWT by applying the formula:

   BLCINCHWT = 1.24408*BLCINCHWT

   to all BLCINCHWT values of units where REGION = '1'

   Enter the sum of BLCINCHWT into the following matrix:

Northeast	23,719,000
Midwest	29,606,000
South	50,679,000
West	28,828,000

15. Do a fourth preliminary adjustment to BLCINCHWT to improve metropolitan status counts.

    a. From published reports, obtain estimated 2013 counts for units by metropolitan status.

Central city	39,980,000
Suburb	64,968,000
Nonmetro	27,884,000

    b. Develop estimates for these same categories using BLCINCHWT with these formulas:

Central city	METRO3 = '1'	37,594,696
Suburb	METRO3 = {'2','3'}	35,667,202
Nonmetro	METRO3 = {'4','5'}	59,570,102

c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

For example, if the estimate in step b for units in central cities is 37,594,696 units, then the ratio for the top cell in step c is 1.06345.

Central city	1.06345
Suburb	0.78178
Nonmetro	1.09061

d. Adjust BLCINCHWT by applying the new adjustment ratios to existing BLCINCHWT using the formulas in step b to determine which BLCINCHWT to adjust by which ratio.

For example, for units in the central cities, create a revised BLCINCHWT by applying the formula:

BLCINCHWT = 1.06345*BLCINCHWT

to all BLCINCHWT values of units where METRO3 = '1'

Enter the sum of BLCINCHWT into the following matrix:

Central city	39,980,000
Suburb	27,884,000
Nonmetro	64,968,000

16. Final adjustment: Adjust the BLCINCHWT to sum to the published totals for 2013 for units by unit type and occupancy status.

a. From published reports, obtain estimated 2013 counts for units by unit type and occupancy status.

The published numbers are:

	Owner-occupied	**Renter-occupied**	**Vacant**	**Seasonal**
Single-family detached	63,414,000	11,714,000	6,771,000	2,425,000
Single-family attached	4,057,000	2,593,000	737,000	228,000
2–4 unit structures	1,390,000	7,760,000	1,414,000	241,000
5–19 unit structures	1,065,000	9,884,000	1,642,000	259,000
20+ unit structures	1,173,000	6,928,000	1,401,000	360,000
Mobile homes	4,577,000	1,340,000	917,000	544,000

b. Develop estimates for these same categories using BLCINCHWT with these formulas:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '1'	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) AND IN13_NUNIT2 = '1'	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '1'	IN13_ISTATUS = {2','3'} AND (8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '1'
Single-family attached	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '2'	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) AND IN13_NUNIT2 = '2'	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '2'	IN13_ISTATUS = {'2','3'} AND (8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '2'
2–4 unit structures	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '3' AND IN13_NUNITS = {2,3,4}	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) IN13_NUNIT2 = '3' AND IN13_NUNITS = {2,3,4}	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) IN13_NUNIT2 = '3' AND IN13_NUNITS = {2,3,4}	IN13_ISTATUS = {'2','3'} AND (8 LE IN13_VACANCY LE 11) IN13_NUNIT2 = '3' AND IN13_NUNITS = {2,3,4}
5–19 unit structures	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '3' AND 5 LE IN13_NUNITS LE 19	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) AND IN13_NUNIT2 = '3' AND 5 LE IN13_NUNITS LE 19	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '3' AND 5 LE IN13_NUNITS LE 19	IN13_ISTATUS = {'2','3'} AND (8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '3' AND 5 LE IN13_NUNITS LE 19
20+ unit structures	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '3' AND 20 LE IN13_NUNITS	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) AND IN13_NUNIT2 = '3' AND 20 LE IN13_NUNITS	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '3' AND 20 LE IN13_NUNITS	IN13_ISTATUS = {'2','3'} AND (8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '3' AND 20 LE IN13_NUNITS
Mobile homes	IN13_ISTATUS = “1” AND IN13_TENURE = 1 AND IN13_NUNIT2 = '4'	IN13_ISTATUS = “1” AND (2 LE IN13_TENURE LE 3) AND IN13_NUNIT2 = '4'	IN13_ISTATUS = {'2','3'} AND NOT(8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '4'	IN13_ISTATUS = {'2','3'} AND (8 LE IN13_VACANCY LE 11) AND IN13_NUNIT2 = '4'

The estimates are:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	60,979,926	11,641,697	7,764,528	2,423,985
Single-family attached	3,595,217	2,323,111	808,273	231,728
2–4 unit structures	1,354,591	7,838,418	1,804,014	263,839
5–19 unit structures	1,032,997	10,306,271	2,267,012	267,528
20+ unit structures	1,061,355	7,349,091	1,921,146	352,712
Mobile homes	4,407,031	1,265,412	1,048,579	523,541

c. Create new adjustment ratios by taking the ratio of the published numbers in step a to the estimates in step b.

   For example, if the estimate in step b for owner-occupied single-family detached units is 60,979,926, then the ratio for the upper left cell in step c is 1.03992.

   The ratios are:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	1.03992	1.00621	0.87204	1.00042
Single-family attached	1.12844	1.11618	0.91182	0.98391
2–4 unit structures	1.02614	0.99000	0.78381	0.91344
5–19 unit structures	1.03098	0.95903	0.72430	0.96812
20+ unit structures	1.10519	0.94270	0.72925	1.02066
Mobile homes	1.03857	1.05894	0.87452	1.03908

d. Adjust BLCINCHWT by applying the new adjustment ratios to existing BLCINCHWT using the formulas in step b to determine which BLCINCHWT to adjust by which ratio.

   For example, for owner-occupied single-family detached units, create a final BLCINCHWT by applying the formula:

   BLCINCHWT = 1.03992*BLCINCHWT

   to all BLCINCHWT values of units where IN13_ISTATUS = '1' AND IN13_TENURE = 1 AND IN13_NUNIT2 = '1'

Enter the sum of BLCINCHWT into the following matrix:

	Owner-occupied	Renter-occupied	Vacant	Seasonal
Single-family detached	63,414,000	11,714,000	6,771,000	2,425,000
Single-family attached	4,057,000	2,593,000	737,000	228,000
2–4 unit structures	1,390,000	7,760,000	1,414,000	241,000
5–19 unit structures	1,065,000	9,884,000	1,642,000	259,000
20+ unit structures	1,173,000	6,928,000	1,401,000	360,000
Mobile homes	4,577,000	1,340,000	917,000	544,000

17. In steps 12, 13, 14, and 15, we successively adjusted the weights to reproduce exactly the count of householders by race, the count of householders by ethnicity, the count of units by region, and the count of units by metropolitan/nonmetropolitan status. The final adjustment reproduces exactly the published counts for occupancy status by structure type. However, each step has disturbed the match achieved by the previous steps so that by the end the only exact match is occupancy status by structure type.

    Now we compare estimates to published counts to see how closely the final weights estimate counts by race, ethnicity, region, and metropolitan/nonmetropolitan status.

Category	Formula	Estimate	Published	Percent error
Region				
Northeast	REGION = '1'	24,076,279	23,719,000	1.5
Midwest	REGION = '2'	28,936,849	29,606,000	-2.3
South	REGION = '3'	50,203,871	50,679,000	-0.9
West	REGION = '4'	29,617,001	28,828,000	2.7
Metropolitan				
Central city	METRO3 = '1'	39,436,398	39,980,000	-1.4
Suburb	METRO3 = {'2','3'}	65,504,986	64,968,000	0.8
Nonmetro	METRO3 = {'4','5'}	27,892,616	27,884,000	0.0
Race/ethnicity				
White only	RACE1 = '01'	93,590,782	93,298,000	0.3
Black only	RACE1 = '02'	14,877,350	15,023,000	-1.0
Hispanic	SPAN1 = '01'	14,741,346	14,681,000	0.4

18. Next we derive estimates of units in public housing and units in HUD-assisted, privately owned properties to see if there are any major discrepancies in this important area. We compare these estimates first to estimates based on AHS weights and then to numbers from the HUD budget.

	CINCH estimates	AHS estimates	Percent error	HUD budget	Percent error
Public housing	994,110	848,769	17.1	1,033,000	-3.8
Vouchers	2,608,772	2,143,564	21.7	2,086,000	25.1
Privately owned	1,413,756	1,170,009	20.8	1,419,000	-0.4

The comparison between estimates based on CINCH weights and AHS weights is the legitimate test. Our weights appear to inflate the count of HUD-assisted units as reported by the AHS. Interestingly the backward-looking CINCH 2013 estimates are much closer to the HUD budget estimates for 2011 for public housing and privately owned assisted housing.

19. As a final check, compare BLCINCHWT to IN13_PWT and IN13_WGT90GEO.

    a. Create two new variables, BLRATPWT and BLRATWGT, as follows:

    IF MXPWT = 0 THEN BLRATPWT = "."

    IF MXPWT GT 0 THEN BLRATPWT = BLCINCHWT/MXPWT

    IF IN13_WGT90GEO = 0 THEN BLRATWGT = "."

    IF IN13_WGT90GEO GT 0 THEN BLRATWGT = BLCINCHWT/ IN13_WGT90GEO

b. Determine means, mode, medians, and key percentiles for the distribution of MXPWT, IN13_WGT90GEO, BLCINCHWT, BLRATPWT, and BLRATWGT.

	MXPWT	IN13_WGT90GEO	BLCINCHWT	Ratio BLCINCHWT/ MXPWT	Ratio BLCINCHWT/ WGT90GEO
Mean	1,750.43	2,152.39	2,496.65	1.46559	1.24927
Mode	2,049.63	2,346.38	3,039.71	1.48306	1.12524
Minimum	8.24	0.00	12.37	0.59664	0.09018
1 percentile	121.18	110.06	171.20	0.87281	0.53811
5 percentile	136.83	174.30	217.21	1.00129	0.71813
10 percentile	446.22	466.37	637.63	1.06310	0.81726
25 percentile	1,537.22	1,519.73	2,069.10	1.21759	1.00228
50 percentile	2,049.63	2,293.70	2,733.44	1.47984	1.22079
75 percentile	2,049.63	2,716.56	3,130.87	1.62598	1.44321
90 percentile	2,148.12	3,250.73	3,417.78	1.93492	1.70999
95 percentile	2,148.12	3,738.79	4,083.96	2.05929	1.88776
99 percentile	3,866.61	5,456.01	4,745.73	2.22783	2.36315
Maximum	26,303.46	29,501.26	51,566.12	2.51252	5.89489

We would expect BLCINCHWT to be larger than either MXPWT or WGT90GEO because we eliminated a large number of cases from the 2013 PUF at step 10. Nothing alarming stands out from this overview of BLCINCHWT. We are satisfied with the ratio of BLCINCHWT to WGT980GEO; its interquartile range is 1.00228 to 1.44321.