Manufactured Housing Survey Public Use File
Technical Documentation

INTRODUCTION

The Manufactured Homes Survey (MHS) is conducted by the U.S. Census Bureau and sponsored by the Department of Housing and Urban Development (HUD). Monthly MHS produces estimates of the average sales price for new manufactured homes placed/sold or intended for sale by geographical region and size of home (Total, Single, Double). Additionally, MHS produces estimates of total shipments on a monthly basis at a national level for homes sold and for-sale for residential use, and nonresidential or other uses. On an annual basis, MHS produces more detailed estimates for placed/sold homes by, size of home, including selected characteristics at a regional level. Average sales price is estimated at the state level. Estimates of average and median square feet at the region level are also produced.

This document describes the Public Use File (PUF) created from data collected monthly for MHS. It provides a brief description of the MHS and the differences between the MHS data and the records available in the PUF. This document also explains how to determine unit-level characteristics from the PUF and how to create estimates and measures of sampling variability from the PUF data. The PUF allows data users greater access to these data than standard tabulations and provides data users the ability to create a wider variety of tables tailored specifically to their data needs.

SAMPLE DESIGN

The monthly sample for MHS is selected from a list of HUD-inspected manufactured home shipments provided to the Census Bureau by the Institute for Building Technology and Safety (IBTS). Due to missing dealer information, not all home sections shipped in a given month may be eligible for sampling. These sections are removed from the frame until the required dealer information is obtained. An adjustment factor is applied to the weighting to account for these cases in the monthly estimates of total homes shipped by status. The list of eligible sections is stratified by dealer region and then a systematic sample of fixed size is taken within each region. A total of 405 home sections are selected monthly and matched to a dealer. Occasionally two sections from the same home are selected and may cause the sample size to be less than 405 when the duplicate is removed. Home sections shipped to FEMA only dealers are selected into the sample with certainty in addition to the 405 sampled each month. Receiving dealers are contacted four months after shipment to determine the placement status and other characteristics of the selected homes.

MHS PUF CONTENTS

The 2017 MHS PUF contains a total 9,520 records, which include a record for each of the 4,859 manufactured homes sampled in 2017 and 4,661 homes shipped to FEMA only dealers.

The 2016 MHS PUF contains a total 7,915 records, which include a record for each of the 4,860 manufactured homes sampled in 2016 and 3,055 homes shipped to FEMA only dealers.

The 2015 MHS PUF contains a total of 5,018 records, which include a record for each of the 4,860 manufactured homes sampled in 2015 and 158 homes shipped to FEMA only dealers.

The 2014 MHS PUF contains a total 4,957 records, which include a record for each of the 4,859 manufactured homes sampled in 2014 and 98 homes shipped to FEMA only dealers. One sampled case in January 2014 was removed since it was a duplicate of another sampled case from that month.

For each record there are six administrative variables and eight variables reflecting relevant responses to MHS questionnaire items.

The form for the MHS can be found at: http://www.census.gov/construction/mhs/form.html

The MHS PUF includes the following information for each record:

- **CONTROL:** An identification variable which includes the shipment month and a random number
- **REGION:** Four Census regions and a fifth code that represents the national level for three or more section homes
- **SHIPMONTH**: Shipment month
- **SECTIONS**: The size of the home with three or more section homes collapsed
- **WEIGHT**: Tabulation weight
- **WGTADJ**: Weight adjustment to account for frame undercoverage

The MHS PUF includes the following respondent-reported variables for each record:

- **STATUS**: The status of the home four months after shipment. Nonresidential placements and other have been collapsed together
- **SQFT**: The square footage of the home. Top and bottom 5% within each region is top coded.
- **PRICE:** (rounded to nearest $1000) this is the actual price if home is placed/sold and the intended sales price if the home is intended for sale. Top and bottom 5% within each region is top coded. Top and bottom codes are rounded to the nearest $100.
- **BEDROOMS:** Number of bedrooms, collapsed to two or fewer and three or more
- Characteristics of sold residential placements: **TITLED**, **LOCATION**, **FOUNDATION**, and **SECURED**

The MHS PUF also contains eight "j" variables (one for each respondent-reported variable) indicating whether a respondent-reported variable is reported (R), imputed (I), or not applicable (9). These variables all have a prefix of "j" in the variable name.

Additional information on all the variables can be found in Appendix A.

NON-RESPONSE AND IMPUTATION

Due to non-response, the MHS PUF contains imputed values for respondent-reported variables, if necessary. Below are the rules used for imputation:

- All records missing a status code have an imputed value for **STATUS**.
- If **STATUS** is "intended for sale" or "placed/sold" then **SQFT**, **BEDROOMS**, and **PRICE** are imputed monthly.

- **TITLED**, **LOCATION**, **FOUNDATION**, and **SECURED** are imputed annually for **STATUS** ="placed/sold".
- All variables except **PRICE** are imputed using a hot-deck method. Collapsed dealer state and home size (single, multi section) are used to define the imputation cells.
- **PRICE** is imputed using a regression model fit to the reported data based on region if more than ten reported records are available in each region; otherwise the model is estimated at the national level. The independent variables in the model are **SQFT** and **SECTIONS** (single, multi section).

PRODUCING ESTIMATES AND THEIR MEASURES OF SAMPLING VARIABILITY FROM THE MHS PUF

Individuals who create estimates using the MHS PUF as input should cite the Census Bureau and the Department of Housing and Urban Development (HUD) as the source of only the input to these estimates. So that weighted estimates and their associated measures of sampling variability may be produced, the MHS PUF contains two weighting components:

1) the inverse probability of selection for a home section divided by the number of sections in the home (WEIGHT)
2) ratio adjustment factor ; $\frac{\#\ homes\ shipped\ in\ month\ m\ of\ size\ i}{weighted\ sum\ of\ sampled\ homes\ in\ month\ m\ of\ size\ i}$ (WGTADJ)
   (Size is either single unit or multi unit)

Estimates created from the PUF are likely to differ from the results of a complete census of all manufactured home shipments. Each monthly sample is one of many probability samples that could have been selected under identical circumstances. Each of the possible samples would yield a different set of results. The standard error associated with a given estimate is a measure of the variation of all the possible sample estimates around the result from a complete census. Low standard errors indicate estimates that are more accurate. In addition to sampling errors, estimates created from the MHS PUF are subject to various response and operational errors: reporting, coding, nonresponse, etc. These nonsampling errors would also occur if a complete census of all shipped homes was conducted under conditions identical to the survey.

To calculate a count estimate of placed/sold homes with a select characteristic, simply sum the product of the sample weight(WEIGHT) and the weight adjustment(WGTADJ) for all homes that have that particular characteristic. Estimates of average sales price or square feet can be determined by multiplying the item value by just the sample weight(WEIGHT), finding the sum and then dividing by the sum of the sample weights (for eligible records, i.e. those that are placed/sold or intended for sale). The variance of the weighted count estimates may be obtained using the following formula:

$$\frac{n}{n-1}\sum_{i=1}^{n}(w_i I - \overline{w_i I})^2$$

Where $I = \begin{cases} 1, & home\ i\ has\ characteristic \\ 0, & home\ i\ does\ not\ have\ characteristic \end{cases}$ and w_i= *WEIGHT$_i$*WGTADJ$_i$*

The variance of the average sales price or square feet estimates can be obtained using the following formula:

$$\frac{n}{n-1}\sum_{i=1}^{n}\frac{[w_i(y_i-\bar{y})]^2}{\sum w_i}\Big/\sum w_i$$

Where $\bar{y}=\sum w_i y_i \Big/ \sum w_i$, y_i is the value of the variable, and w_i= *WEIGHT*$_i$

CONFIDENTIALITY AND DISCLOSURE AVOIDANCE TECHNIQUES

In accordance with federal law governing census reports (Title 13 of the United States Code), no data are published that would disclose the identity of an individual manufactured home. This applies to both the records on the MHS PUF and tabulations produced from the PUF. The Census Bureau's Internal Disclosure Review Board sets the confidentiality rules for all data releases. A checklist approach is used to ensure that all potential risks to the confidentiality of the data are considered and addressed.

Methods for preventing the disclosure of reported data in the MHS PUF include an extensive review and analysis of internal tabulations produced from the PUF. The internal tabulations were used to determine the level of detail that could be shown for variables included in the MHS PUF. All homes, with the exception of larger homes comprised of three or more sections, include geographic detail at the region level. Top-/bottom-coding of continuous variables square feet and price and collapsing of some levels of characteristic variables were the methods used to prevent data disclosures.

Appendix A: MHS PUF File Contents

Variable Name	Variable Description	Values/ Format	Details	Notes
CONTROL	Random ID # for home	YYMMNNNN		
REGION	Region (dealer region if not placed; otherwise, placement region)	1,2,3,4,5	1= Northeast 2=Midwest 3=South 4=West 5=United States	US level "5" only used for 3 or more section homes
SHIPMONTH	Shipment month	YYYYMM	YYYY=year, MM=month	
SECTIONS	Size of home	1,2,3	1= Single 2=Double 3=3 or more sections	
WEIGHT	Sample weight	XX.XXXX	Probability of selection/ # of sections	
WGTADJ	Weight adjustment	XX.XXXX	Ratio adjustment of sample weight	Use for estimation of shipment totals
STATUS	Status of home four months after shipment	1,2,3	1= Intended for Sale for Residential Use 2=Placed/Sold for Residential Use 3=Non-res/Other	
PRICE	Sales price or intended sales price	$XX,X00	Rounded to nearest $1000 (unless its topcoded and then rounded to nearest $100) 9 is for Nonapplicable cases	Top and bottom 5% are coded to the average of those records to preserve distribution
SQFT	Square footage of home	X,X00	Rounded to nearest $100 9 is for Nonapplicable cases	Top and bottom 5% are coded to the average of those records to preserve distribution
BEDROOMS	Number of bedrooms	1,3,9	1= 2 or less 3= 3 or more	9 is for Nonapplicable cases
TITLED	How a home is titled	1,2,3,9	1= Real Estate 2= Personal Prop 3= Not titled	9 is for Nonapplicable cases

LOCATION	Where a home is placed	1,3,9	1 = Inside manuf. home communities 3 = Outside manuf. home communities	9 is for Nonapplicable cases
FOUNDATION	Type of foundation for home, if any.	1,2,3,9	1= Masonry/Concrete 2=Steel or Other 3=Blocks	9 is for Nonapplicable cases
SECURED	How a home is secured (if not on permanent masonry/concrete)	1,3,9	1= Tie down straps or other 3=Not secured	9 is for Nonapplicable cases
JSTATUS	Imputed indicator for status code	R,I	R= Reported I=Imputed	
JPRICE	Imputed indicator for price	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JSQFT	Imputed indicator for square feet	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JBEDROOM	Imputed indicator for bedrooms	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JTITLE	Imputed indicator for titled	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JLOCATION	Imputed indicator for location	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JFOUNDATION	Imputed indicator for foundation	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases
JSECURED	Imputed indicator for secured	R,I,9	R= Reported I=Imputed	9 is for Nonapplicable cases