

Description of Methodology to Benchmark Existing Home Sales

In 2004, the NATIONAL ASSOCIATION OF REALTORS® Economics and Research Division took on the task of rebenchmarking its existing home sales data for the fifty states and the District of Columbia to the 2000 Census. Along with the rebenchmarking, NAR cleaned up its sample database for the EHS series. NAR therefore believes it now has the most reliable possible estimates of existing home sales for the Nation and the four Census regions, as well as for the 50 States and the District of Columbia.

Below we set forth the steps in the benchmarking and existing home sales estimation process. They include:

- The Determination of a Benchmark for Single-family Existing Owner-Occupied Home Sales

- The Determination of a Benchmark for Single-family Existing Home Investor and Vacant Sales

- The Determination of a Benchmark for Existing Condominium Sales

- Using the Benchmark to Get Quarterly Totals

The NATIONAL ASSOCIATION OF REALTORS® existing home sales data are estimates arising from a sample of Multiple Listing Service (MLS) sales. As time changes, so does the relationship between any sample and the universe the sample is representing.

The EHS is not alone in its need for rebenchmarking on occasion. Data that the government reports on a regular basis, such as Gross Domestic Product and the Consumer Price Index, are also based on survey data. The government thus rebenchmarks these data on an occasional basis.

We note that NAR took on this rebenchmarking project on its own initiative, so that it will continue to provide analysts and the public the most reliable data possible on the housing market.

II. The Determination of a Benchmark for Single-family Existing Owner-Occupied Home Sales

The benchmark for single-family existing home sales comes in a fairly straightforward manner from the Public Use Micro-sample (PUMS) of the 2000 U.S. Census. Specifically, we have used the "PUMSA" sample, or the 5 percent sample. This sample gives us detailed information on one in twenty housing units enumerated in the 2000 Decennial Census. Because the census draws a stratified (i.e., non-random) sample, each household in the sample receives a weight that reflects how representative it is of the overall population. Altogether, the five percent sample allows us to draw inferences about housing characteristics for the nation based on over 5,000,000 households.

For every household in the sample we determine the following:

- whether it lives in an owner-occupied house

- whether it lives in a house that it moved into during or after 1999
- whether it lives in a single-family attached or detached house
- whether it lives in a house that was built before 1999

Houses that meet all the above criteria are Single-family Owner-Occupied Existing Home Sales for 1999 and the first three months of 2000 (the census does not ask about a particular year). Note what is not included: Mobile Homes, New Homes, Multifamily Homes, and Commercial Properties.

The great benefit of using the PUMS data set is that it is so comprehensive and large. However, when the Census asks respondents when they moved into their homes, it does not ask about a particular calendar year, but rather whether the respondent moved within the fifteen-month period described above. Because we are interested in getting annual sales rates, we must convert these fifteen-month sales numbers into annual numbers.

A straightforward way for doing this would be to multiply the fifteen-month number by .8 to get an annualized number. The problem with this approach is that housing sales are very seasonal: generally speaking, not as many houses sell in the first three months of the year as in the other three month periods. Therefore, for each state, we take the sum of the raw volume for the four quarters of 1999 and divide this by the sum of the raw volume for the first four quarters of 1999 plus the first quarter of 2000. Multiplying this quotient by the fifteen-month count allows us to scale the 15-month period to a 12-month period, and gives us our annual benchmarks for single-family existing owner-occupied home sales for 1999 for the 50 states and the District of Columbia.

We exclude mobile homes, new homes, multifamily homes, and commercial properties from our total because we do not intend to capture these sales in our survey of MLSs. All of our respondents do, however, report single-family MLS sales, including single-family renter-occupied (or investor) and single-family vacant homes. The total calculated above does not include sales of these homes, so we must add these to our benchmark for single-family existing owner-occupied homes sales to get a benchmark for single-family existing home sales. We describe how we do this next.

III. The Determination of a Benchmark for Single-family Existing Home Investor and Vacant Sales

In order to determine the number of investor sales in 1999, we rely on information from the U.S. Census' 2001 Residential Finance Survey. This micro-data set contains approximately 68,000 observations. We reduce the data set until we have observations that meet the following criteria:

- A single-family detached or attached house
- A house that is rented or vacant

- A house that already existed in 1999

Having reduced the number of observations to those that meet the above criteria, we then examine how many of these houses transacted in 1999. We include only existing home transactions—we exclude sales from builders to investors. The number of houses within this data set that transacted divided by the number of observations in the data set that meet the three criteria described above give us a single-family existing home investor and vacant sales turnover ratio. The procedure so far described produces a national turnover ratio.

Altogether, we create fourteen turnover ratios: one for the four Census regions: Northeast, Midwest, South, & West, and one for these ten States: Massachusetts, New Jersey, New York, Pennsylvania, Florida, Texas, Virginia, Ohio, Illinois, and California. Unfortunately, the sample size of the Residential Finance Survey is too small to determine reliable turnover ratios for states other than the ones mentioned above. We therefore apply regional turnover ratios when determining the benchmark for investor and vacant sales in these places.

The final step is to multiply the appropriate turnover ratio by the stock of single-family investor and vacant homes for a particular location: this will give us an estimate of single-family investor and vacant sales. The stock number comes from tabulations of the 2000 Public Use Micro Sample of the U.S. Census. The PUMS allow us to determine the share of single-family houses that were either rented or vacant and that were built before 1999 (i.e., that were not new). We multiply this share by the total number of single-family homes enumerated by the complete 2000 census to get the size of the single-family existing housing investor and vacant stock.

To give an example, we use the PUMS data and the Census enumeration to calculate that California had a total of 1,988,475 existing single-family vacant and detached units in 1999. We use the RFS data to calculate that the turnover rate for housing of this type in California for that year was 6.5 percent. Multiplying 1,988,475 by .065, we estimate that there were 129,250 sales of vacant and investor single-family existing homes in California in 1999.

IV. The Determination of a Benchmark for Existing Condominium and Cooperative Sales

The procedure for determining existing condominium and cooperative sales is similar to the procedure for existing single-family home sales. For owner-occupied condominium sales, we once again rely on the five percent PUMS data, and find observations that meet the following criteria:

- household responds that it paid a condo/coop fee
- household responds that it moved in 1999 or 2000
- household responds that unit was built before 1999
- household responds that it does not live in a single-family attached or detached unit (so we do not double count units that are both single-family and condominiums).

For investor and vacant condo sales, we follow the exact same procedure as that for investor and vacant single-family sales, with two exceptions. We look only at the sample of condominiums that are not single-family sales in the Residential Finance Survey, and we only determine regional turnover rates (rather than regional rates plus the rates in the four largest states). The reason for the latter restriction is a practical one—there are not enough observations involving investor and vacant condominium sales for the four individual states to get reliable turnover rates for condominiums in these states.

V. Refinements to Methodology

Following NAR’s philosophy of continually improving its benchmark, the Association refined its 1999 methodology. We give a brief overview of the refinements here.

Refinement #1: The core of our benchmark is the following question from the Census Survey: “*When did this person move into this house, apartment, or mobile home?*”

In our previous methodology, we made the assumption that respondents who said they *moved* into an owner-occupied unit, also *purchased* the unit that same year. And indeed, this is true for the vast majority of such households. However, analysis of American Housing Survey data shows that a number of households who moved into a housing unit during a particular year purchased the unit in any earlier year. In estimating our benchmark year, we are only interested in counting home sales transactions that occurred in 1999. We must therefore remove households who moved during the benchmark period but purchased in an earlier period in order to capture accurately the number of existing home sales that occurred during the benchmark period. To make this adjustment, we use the 1999 AHS data to estimate the percentage of households who moved into the household in 1999, but did not purchase the property in that same year. We then remove these from the total, giving us a more accurate estimate of *actual* home sale transactions in 1999.

Refinement #2: NAR does not intend to capture new home sales in its survey of MLSs. However, there are a percentage of new homes that are bought and resold in the same year. When a new home is resold, it typically involves a real estate professional, and it is recorded into an MLS as an existing home sale. Thus, these “quick” sales of new homes are picked up in NAR’s sample as existing home sales. Since we are mapping our benchmarks to our sample of MLSs, it is therefore necessary to adjust our benchmarks to include these “quick” sales. Given limitations in using Government data, the best available information for estimating this portion of sales came from NAR’s survey of Home Buyers & Sellers. From this survey, we were able to estimate the percentage of new homebuyers who owned their home for “one year or less”, and we made adjustments to the benchmark based on these results.

Refinement #3: In 1999, we did not include Census data on single-family homes on 10 acres or more. These statistics were excluded from the universe because, at the time, it was common practice to define these properties as farms and thus commercial and not residential. Over time, however, these larger properties have been entering the MLSs as residential sales. Through

improved data collection and electronic uploads, we can now identify these sales as existing home sales, and we therefore now include them in the benchmarks.

VI. Conclusion

In addition to rebenchmarking the existing home sales series, the NATIONAL ASSOCIATION OF REALTORS® also added a few more Realtor Boards & MLSs to create a stronger sample in revising Existing Home Sales data forward from 1999 to the present. We decided to include boards we believe should be part of the Existing Home Sales panel for various reasons; for example, we include boards in areas that have experienced significant growth within the past 10 years, boards in underrepresented areas, boards with terrific history of participation, etc. As a result, NAR's sample of home sales has been vastly improved to more accurately represent national/regional and state level housing activity.

Because the basis for NAR's database of existing home sales volume is a sample of sales throughout the country, a mechanism is necessary to "blow up" the sample sizes to reflect the nation as a whole. Fortunately, the decennial Census and the Residential Finance Survey provide as comprehensive of a snapshot of the country as is available anywhere. Therefore, while a few simplifying assumptions have been made to calculate our volume benchmark, they do a good job of providing a basis for adjusting our regularly collected information on home sales nationwide.

Please direct your questions and comments to:

Richard K. Green, drgreen@gwu.edu, Professor of Real Estate Finance, The George Washington University, Washington, DC 20052

Kevin J. Thorpe, kthorpe@realtors.org, Manager of Home Sales Surveys
NATIONAL ASSOCIATION OF REALTORS®