

ON THE HAZARDS OF INFERRING HOUSING PRICE TRENDS USING MEAN/MEDIAN PRICES

Chee W. Chow, Charles W. Lamden School of Accountancy, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, chow@mail.sdsu.edu

Damon M. Fleming, Charles W. Lamden School of Accountancy, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, dfleming@mail.sdsu.edu

James E. Williamson, Charles W. Lamden School of Accountancy, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, james.williamson@mail.sdsu.edu

ABSTRACT

Overall conclusions about the real estate market are generally based on the trend in mean and/or median prices. The objective of this article is to point out that mean/median statistics can lead to completely wrong conclusions about where prices are headed. We first use a simple example to show how using mean/median prices as a barometer can result in erroneous conclusions about the state of real estate prices. Then we illustrate how some recently reported statistics about house prices could have come about from either a rising or falling market. Finally, we suggest some ways to improve the statistics used as a reference.

Key words: Real estate; average home prices; statistics.

INTRODUCTION

Most people consider owning one's home an integral part of the American dream. This mentality, in conjunction with the easy availability of credit, undoubtedly caused many people to rush into the real estate market in the past few years. The burst of that real estate bubble has left countless families stranded with the prospects of foreclosure and greatly diminished personal wealth. Governments from the local to federal levels also have been pushed into harsh economic straits. The state of California, for example, is contending with a multi-billion dollar deficit, due in no small part to markedly lower real estate taxes. These, in turn, had resulted from shriveled real estate values and greatly reduced real estate transaction volumes.

Now, signs are beginning to appear that the economy has reached bottom and started to recover. In the real estate sector, stories that prices have stabilized and perhaps even improving slightly are beginning to appear in the popular press. Are these observations about real estate price trends—be they favorable or unfavorable—justified? This is an important question because if people believe that house prices are on their way up again, this could re-ignite their interest in real estate investments. To a large extent, it was the perception of rapidly escalating real estate prices that had helped to fuel the last real estate bubble, with obviously disastrous results.

News stories about the real estate market generally contain stories of human interest (e.g., what some families had experienced) and other details. However, overall conclusions invariably are based on the trend in mean and/or median prices (the latter is the price where, if you lined up all of the prices from the highest to the lowest, 50 percent of the prices would be above it, and the remaining 50 percent would be below it). The objective of this article is to point out that mean/median statistics can lead to

completely wrong conclusions about where prices are headed. The entire housing market can be nose-diving at the same time that mean/median house prices show an upward trend. Conversely, a rising real estate market can be accompanied by lower and lower mean/median price statistics.

In the rest of this article, we first use a simple example to show how using mean/median prices as a barometer can result in erroneous conclusions about the state of real estate prices. Then we illustrate how some recently reported statistics about house prices could have come about from either a rising or falling market. Finally, we suggest some ways to improve the statistics used as a reference.

A SIMPLE EXAMPLE

In Home Town, USA, there are 100 homes, of which 20 are mansions and 80 are cottages. In 2009, 6 mansions were sold for \$500,000 each and 5 cottages were sold at \$200,000 each. The local newspaper reported that the median price of homes sold in 2009 was \$500,000 and that the mean (average) price of homes sold in 2009 was \$363,636.

In 2010, 15 homes were sold and there actually was a 10 percent increase in the prices of all homes in Home Town. Of the 15 homes sold, 3 were mansions priced at \$550,000 each and 12 were cottages priced at \$220,000 each. The local newspaper reported that the median price of homes sold in 2010 was \$220,000 (a decrease of \$280,000) and that the mean price of homes sold in 2010 was \$286,000 (a decrease of \$77,636). Although the median and mean values for each year were quite different, both statistics indicated a substantial decrease from 2009 to 2010. Since the prices of all homes had gone up 10 percent between the two years, the picture painted by the trends in mean/median prices is obviously wrong.

What caused these statistics to produce a picture that is opposite to the underlying reality? The basic reason is that both of them are “driven” by two factors: the mix of houses actually sold and their prices. If one or both of these factors are not representative of the population of homes in the community, then the mean/median prices of houses sold would not accurately reflect the underlying state of affairs.

To see this, first consider the mean house prices for the two years. For 2009, the mean price was weighed more heavily towards the mansions (6 out of 11 homes sold) than cottages. Even though both mansions and cottages had sold for lower prices in 2009 than in 2010, the mean price for 2010 was weighted much more heavily towards cottages because they constituted the bulk of house sales in 2010 (12 out of 15). Since cottages are much lower priced than mansions, this shift in relative weights between the two years caused the mean price for 2010 to be lower than that in 2009, even though both mansions and cottages had increased in price.

In addition to being affected by the mix of houses sold, the mean also is highly dependent on their prices. To see this, assume that in 2010, two of the three mansions sold had gone for \$1 million apiece. This would have inflated the mean price for 2010 to \$346,000 without affecting the value of the median. Or, assume that in 2010, 4 of the cottages had sold for \$100,000 each instead of \$220,000. This would have caused the mean price for 2010 to drop to \$254,000, again without disturbing the median price. It is because the mean can be drastically affected by a few extreme values (while medians tend to be immune) that popular reports tend to focus on the median. Hence, to keep things simple we will limit our subsequent discussion to the median.

For 2009, because more mansions were sold than cottages, the price that is in the middle is the price of a mansion (\$500,000). For 2010, with cottages making up the bulk of houses sold, the median price is that of a cottage (\$220,000). This is why median prices dropped between the two years. If the mix of houses sold in either year had been different, a different picture could easily have emerged. For example, suppose that only 4 mansions had been sold in 2009 as compared to 5 cottages. In that case, the median price for 2009 would have been the price of a cottage (\$200,000), and comparing the median prices between the two years would have revealed an upward trend (from \$200,000 to \$220,000). While an

upward trend obviously is preferred to one heading downwards, the key is to get an accurate picture of where the entire real estate market is headed. And as this simple example has illustrated, the median can be a highly flawed barometer.

AN EXAMPLE BASED ON REPORTED STATISTICS

On Sunday, October 11, 2009, an article in the San Diego Union Tribune (SDUT) reported that the national median price of homes sold in October, 2008, was \$213,000 while the national median price of homes sold in August, 2009, was \$195,000 [1]. With no accompanying qualifying information, the reader might easily be led to believe that house prices had fallen another 8 percent in the past year. However, also on October 11, 2009, Parade Magazine, which was included in the SDUT package, reported that the size of new houses had shrunk last year and that new houses under construction through June were nearly 200 square feet smaller than the average new home size of 2500 feet in 2007, a drop of 8 percent[2]. This decrease in house size could have accounted, at least in part, for the drop in median prices from the earlier period. This observation alone cautions against looking only at price statistics without considering characteristics of the houses that were sold.

But if there has not been a change in house size or other characteristics, could one then draw conclusions about the state of the real estate market based on the trend in median prices? At the California Association of Realtors annual conference in San Jose, California on October 7, 2009, it was reported that, "Prices rose to a record median \$560,300 in 2007 before falling more than 50 percent to this year's estimated \$271,000" [3]. The median prices were taken from 346,900 reported sales in 2007 and an estimated 540,000 reported sales in 2009. Are we to conclude that house prices in California had dropped more than 50 percent? The answer is: "It is impossible to tell." In fact, looking at the median price statistics alone, one cannot even rule out that the prices of houses had increased from 2007 to 2009.

To see this, assume that there are three types of houses in California: mansions (highest priced), lodges (medium priced) and cottages (lowest priced). In 2007, a total of 347 (in thousands) houses were sold in California. Of these, assume that 100 were mansions priced at \$1,000,000 each, 150 were lodges sold at \$560,300 each, and the remaining 93 units were cottages that had sold for \$135,500 each. Based on these (fictitious) figures, the median selling price for 2007 would be \$560,300, the price of a lodge. Now consider the following three scenarios:

Scenario A

In 2009, 100 mansions were sold for \$900,000 each (a price decline from 2007 of 10 percent); 200 lodges were sold for \$271,000 each (a major price decline of over 51 percent), and 240 cottages were sold for \$121,950 each (a price decline of 10 percent). Thus, house prices had declined across the board. However, the declines for mansions and cottages were only moderate, and only the lodges had suffered a major setback. But because most of the units sold were lodges, the median price would be that of a lodge, and it would paint an overall bleak picture.

Scenario B

In 2009, all houses had a price increase of 50%. 50 mansions were sold for \$1,500,000 each; 100 lodges were sold for \$1,120,600 each, and 390 cottages were sold for \$271,000 each. In this case, because relatively few medium and higher priced units were sold, the median price is that of a cottage: \$271,000. Thus, while the median price for 2009 was over 50 percent lower as compared to the year 2007, the

underlying reality is that all house prices had gone up by 50 percent rather than declined by about such a percentage.

Scenario C

In 2009, the prices of both mansions and lodges declined by 10 percent whilst the prices of cottages doubled. 50 mansions were sold for \$900,000 each, 50 lodges were sold for \$504,270 each, and 440 cottages were sold for \$271,000 per unit. Again, because of the volume of cottage sales, the median price is that of a cottage at \$271,000.

One could easily come up with other combinations of sales mix and prices that could have produced the reported median prices for 2007 and 2009. These combinations could include price increases for some types of dwellings and price decreases for others, price decreases for all, but to different degrees, or even price increases of equal or unequal extents across dwellings. Thus the main point of these illustrations should be clear: the median is the outcome of interactions among underlying factors that tend not to be reported along with the median statistic. Because the same value of the median can be produced by widely diverse sets of underlying factors, using the median value to draw inferences about the general direction of housing prices can be perilous. Median prices can be trending up when all houses are suffering a price decline, holding steady, or increasing, but at a different rate than that of the change in median values. Conversely, median prices can be declining at the same time that all house values are on the rise.

HOW TO GET A BETTER PICTURE

Since the problem with the median is that it may not reflect the underlying population of housing units and their prices, one might suggest coming up with a weighted average price based on the actual mix of houses in the area of interest, rather than just those that are sold. While this approach seems reasonable, it is unlikely to be feasible. Coming up with the weighted average not only requires considering a large number of individual units, but there also is the issue of estimating each unit's current value. With houses having many features (location, age, size, amenities, view, quality of construction, etc.) it would be costly, if not impossible, to form such estimates. The challenges of factoring in such differences also limit the usefulness of oft-reported statistics like mean or median price per square foot. While converting house prices to a per square foot basis does have the advantage of controlling for differences in house sizes, this statistic still is influenced by differences in the many other house characteristics. Furthermore, the price per square foot for the houses actually sold still would be subject to the limitations we had illustrated earlier.

We believe that a workable option may be to come up with a number of prototypical houses to represent major segments of the housing units in a community of interest. This can be done by making use of information typically available in real estate listings. For example, one class of houses might be from 1,800 to 2,200 square feet in size, with four bedrooms, 2-1/2 to three bathrooms, and in an "average" neighborhood. Another might be 1,200 to 1,500 square feet in size, with three bedrooms, up to two bathrooms and in a lower income neighborhood. Comparing median prices for a given type of house across time can free the analysis from the impacts of changing sales mixes as well as those of different house features. And the more elaborate the house classifications (e.g., also considering age, proximity to shopping, ratings of local schools), the more the price statistics will accurately reflect the current state and direction of the real estate market.

SUMMARY

Median and mean house prices often are relied upon to make inferences about the current state and trend in housing prices. This article has demonstrated that these statistics can be highly flawed bases for making such inferences. In particular, examples are used to show that the trend of median prices can over or under represent changes in housing prices, and even move in an opposite direction to the underlying reality. We also suggest a way to improve the usefulness of statistics for making such assessments.

REFERENCES

- [1] T. Paradis. Numbers tell volatile financial story: October marks crash, record high for stock market, *The San Diego Union Tribune*, October 11, 2009, A 3.
- [2] Anonymous, American Homes Are Getting Smaller, *Parade Magazine*, October, 11, 2009, 6.
- [3] R. Showley. Decline Seen in Home Sales, *The San Diego Union Tribune*, October 8, 2009, C 4.