

ASU-Repeat Sales Index (RSI) Methodology

Karl L. Guntermann
Taylor Professor of Real Estate

Alex Horenstein
Research Associate

Introduction

The Arizona State University Repeat Sales Indices (ASU-RSI) are constructed using an approximated algorithm to the Robust Interval and Value Weighted Arithmetic Repeat Sales algorithm (Robust IVW-ARS) used by S&P/Case-Shiller Home Price Indices¹ (S&P/C-S). We tried to approximate the ASU-RSI algorithm and its properties as much as possible to the one used by S&P/C-S². The cleaning process used on the S&P/C-S data is proprietary so any differences between the two indices would be due to differences in the cleaning process used with the data, possible differences in the 3-step procedure used (more specifically in the estimated variance) and because of differences in the data used for estimating the index. One difference between the databases is that the ASU-RSI includes FSBOs (for sale by owner) whereas the S&P/Case-Shiller index uses Multiple Listing Service (MLS) data. Graph 1 shows the ASU-RSI for the Metro Area and the S&P/Case-Shiller index for Phoenix from January 1989 to October 2006.

Data included

Because the ASU-RSI is calculated for regions and selected cities, the cleaning process, which is intended to eliminate data errors, reflects geographic differences in house prices throughout the metropolitan area. Only single-family houses are included in the calculation of the indices. In order for a repeat sale to be included in the estimation of an index it should have the following characteristics:

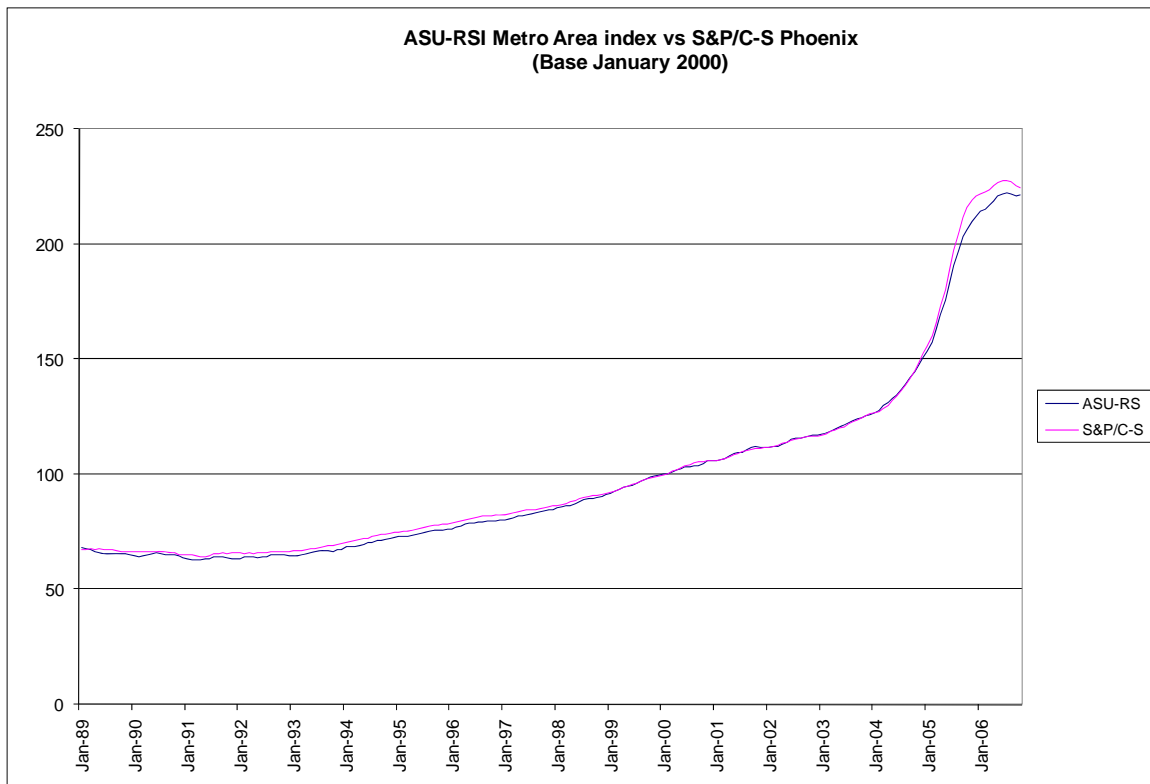
- All sales must be between unrelated parties.
- It must not be the sale of a new house.
- The period between sales should be at least six months.
- The price of a house should be less than \$15 million.

¹ For a more detailed survey refer to the “S&P/Case-Shiller Home Price Indices” index methodology at: http://www2.standardandpoors.com/spf/pdf/index/SPCS_MetroArea_HomePrices_Methodology.pdf

² Because of proprietary rights we don't know the exact algorithm used to construct the S&P/Case-Shiller Home Price Indices.

- House price per square foot must be within the range of \$10 to \$800.
- House prices below \$10,000 after the year 2000 are dropped except for Phoenix, Scottsdale, Paradise Valley and Fountain Hills where the minimum price after 2000 is \$30,000.
- Annual appreciation may not exceed 80 % while annual depreciation can be no greater than 60% per year.

Graph 1



ASU-RSI methodology

The ASU-RSI indices are calculated in two stages, a pre-base estimation and a post-base estimation.

In the first stage (pre-base estimation), the Robust IVW-ARS is estimated using a 3-step methodology for the pre-base period. This consists in estimating all index points from January 1989 until the end of the base period simultaneously. Table 1 shows the length of the base period for each index. The base period for the Metro Area is the same as the

S&P/Case-Shiller index. In this stage we also calculate the variances of the mispricing errors that will be used to correct for heteroskedasticity. Sometimes, in order to achieve a more consistent estimator of the mispricing errors, the period used for estimating the variance is longer than the base period. For the results on the estimated variances and the periods used please see the next section.

Table 1

INDEX	BASE PERIOD
ASU-RSI Metro Area	January 2000
ASU-RSI Phoenix & Central Region	January 2000
ASU-RSI Northeast Region	January 2000
ASU-RSI Southeast Region	January 2000
ASU-RSI Northwest Region	January 2000
ASU-RSI Southwest Region	January 2007
ASU-RSI Chandler	January 2007
ASU-RSI Glendale	January 2007
ASU-RSI Mesa	January 2000
ASU-RSI Peoria	January 2007
ASU-RSI Scottsdale-Paradise Valley	January 2000
ASU-RSI Sun City / Sun City West	January 2007
ASU-RSI Tempe	January 2007

The second stage consists of updating the ASU-RSI each month by adding the new values to the existing indices. As is done with S&P/C-S, the indices are estimated using a chain-weighted procedure where the new index is conditional on all the previous values of the index but not on future values. This method is used to limit revisions of the estimated indices.

Three-month moving average: Each repeat sale is used for the month when it occurs and for the next two months (keeping the interval between sales constant). For example, a repeat sale that occurs in November 2006 is included in calculations for November 2006, December 2006 and January 2007 and is part of the November 2006 index value. The use of a three-month moving average is consistent with S&P/Case-Shiller.

1st Stage (Pre-Base simultaneous Estimation)

A 3-step estimation procedure is used at this stage. The first and second steps are the same first two steps as for the Interval Geometric Repeat Sales index (I-GRS) in Shiller (1991) or Case and Shiller (1987). These steps are used to calculate the covariance matrix that will be used in the 3rd step to correct for heteroskedasticity and obtain the final index and they will also be used in the second stage. As explained in Shiller (1991), the same matrix for the generalized least square process estimated for the I-GRS can be used for the IVW-ARS. Thus, in our 3rd step we calculate the entire index at once using the IVW-ARS method that can be described as follows:

$\hat{\beta} = (Z' \hat{\Omega}^{-1} X)^{-1} Z' \hat{\Omega}^{-1} Y$ where $\hat{\beta}, \hat{\Omega}, Z$ and X are as described in Shiller (1991), page 123.

It is known that when there are observations with more than 1 repeat sale the matrix Ω is not diagonal but block-diagonal with each block corresponding to an individual house. Nevertheless, the difference in the estimation between using this block-diagonal matrix and a diagonal one (as if each repeat sale corresponds to a different house) is very small and the ASU-RSI index uses a diagonal $\hat{\Omega}$.

2nd Stage (Post-Base chain-weighted Estimation)

In this stage the ASU-RSI index is calculated using a three-month moving average³ and the estimated variance calculated in stage one⁴. Each index is calculated conditional on all the previous values of the index and independent of all future values.

Estimated Variances for the mispricing errors

Given the calculation of the I-GRS index (first step of the Pre-Base estimation), the predicted house price in period t given the price in period s is:

³ See Standard & Poors (May 2006). “S&P/Case-Shiller Home Price Indices index methodology”

⁴ Unless a revision of the index is made

$$\ln(\hat{P}_{it}) = \ln(P_{is}) + (\hat{\beta}_t - \hat{\beta}_s)$$

where the $\hat{\beta}$ represents the logarithm of the estimated value of the index using the I-GRS methodology⁵. Then, as in Case and Shiller (1987) it is assumed that the error term of the GRS estimator has a variance equal to $2\sigma_N^2 + \sigma_H^2 D_i$ where D_i is the interval between the sales.

In practice, sometimes the estimation of σ_N^2 can be negative. That happened while calculating some of the ASU-RSI indices. However, when $\hat{\sigma}_N^2$ was negative, the estimator was not significantly different from zero at a 95% level of confidence and was collapsed to zero. Table 2 shows the estimated value of these variances for the different estimated indices. It also shows the length of the period that was used to estimate the mispricing error.

Revisions

The chain-weighted procedure is used to accurately estimate the trend of the index while minimizing the revisions to it. Ideally, every time the indices are updated, all repeat sales for the new month would be included. However, because of delays in the collection and reporting processes, some repeat sales may not be included in the initial calculation. If additional sales for the new month are reported within the next two months, they would be included in the later calculations but sales reported after the indices are updated would be omitted. Better data or an improved methodology might lead to future revisions of the indices.

⁵ See Calhoun, C.A. (1996), page 8.

Table 2

	$2\sigma_N^2$	σ_H^2	R2	Observations
Metro Area *	0.0027558 (1.88)	0.000853532 (29.74)	0.0085	188243
Central / Phoenix **	0.013181 (8.79)	0.00079803 (37.64)	0.0115	205855
Northeast **	0.0092178 (4.57)	0.00066598 (21.33)	0.0090	84156
Southeast **	0.0041873 (3.36)	0.00048072 (26.01)	0.0065	175526
Northwest **	0.0066233 (3.98)	0.00050836 (20.99)	0.007	106883
Southwest **	-0.00804 (-1.68)	0.0010883 (13.41)	0.0267	11599
Chandler ***	-0.0017173 (-0.75)	0.00049924 (13.62)		50669
Glendale ***	0.0026139 (0.95)	0.00056776 (14.40)	0.0068	51193
Mesa *	-0.0036297 (-0.92)	0.00076235 (10.27)	0.0078	24594
Peoria ***	0.0031587 (0.99)	0.00041033 (8.43)	0.0056	21773
Scottsdale - Paradise Valley *	-0.000819 (-0.21)	0.00090044 (11.65)	0.0091	26687
Sun City - Sun City West ***	0.016075 (5.52)	0.00030255 (7.47)	0.0046	21202
Tempe ***	0.0053715 (1.52)	0.00040844 (8.34)	0.0048	24669
t-statistics are in parenthesis				
* January 1989 - January 2000				
** January 1989 - June 2007				
*** January 1989 - January 2007				

Bibliography

Calhoun, C.A. (1996). "OFHEO House Price Indexes: HPI technical description", *Office of Federal Housing Enterprise Oversight*.

Case, K.E. and Shiller, R.J. (1987). "Prices of Single Family Homes since 1970: New Indexes for Four Cities", *New England Economic Review*, 45-56.

Shiller, R.J. (1991). "Arithmetic Repeat Sales Price Estimator", *Journal of Housing Economics*, 110-126.

Standard & Poors (May 2006). "S&P/Case-Shiller Home Price Indices index methodology".