A COINTEGRATION TEST TO VERIFY THE HOUSING BUBBLE

Bala Arshanapalli, Indiana University Northwest
William Nelson, Indiana University Northwest

ABSTRACT

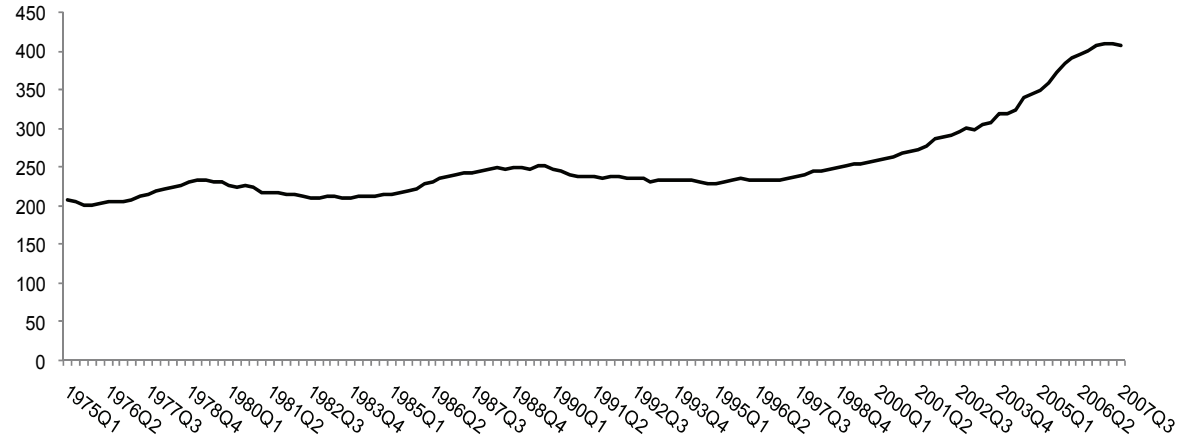
Housing prices in the US rose rapidly from 2000-2007Q3. Based on this evidence, the financial and general press concluded the US experienced a housing bubble. The efficient market theory denies the possibility of a bubble. This paper applies the statistical technique of cointegration to substantiate the presence of a housing bubble. The paper finds the statistical evidence consistent with the presence of a housing bubble in the period 2000-2007Q3 and not the underlying economic conditions.

JEL: C53, G12, G18

INTRODUCTION

Recently housing prices in the US have risen substantially. Figure 1 shows the trend in the US Housing index from 1975 to 2007Q3. Clearly, from 2001 the US has experienced several years of strong home price increases. This sort of evidence has led the financial and the general press to conclude the US housing market has experienced a bubble. Today the press speculates that this bubble has deflated and may soon pop if it has not already popped. Nevertheless, prolonged rapid increases in prices do not imply the presence of a bubble. Nor does an ensuing sharp price decline following a price run-up necessarily imply a bursting bubble. The former event may simply reflect the changes in the fundamental economic factors, such as the low level of mortgage rates in the case of housing prices. Lowenstein (2007) explores this possibility in the general press. The latter event may be a result of a sudden reversal in the underlying fundamental economic factors.

Figure 1: Home Prices in the United States (1975-2007Q3)



US Housing index from 1975 to 2007Q3

The bursting of a real estate bubble has important implications for the US economy. Residential real estate is an important component of householder wealth. In 1996, it represented 39% of household wealth; today it represents 49% (Roubini, 2006). Householder wealth is supporting the high current level of consumption. If a significant portion of real estate wealth disappears, it is likely that consumption will

fall considerably. It is important to remember that the housing boom was strongly encouraged by the low interest during the 1990 decade. Many mortgages are adjustable rate. In today's environment of higher interest rates the higher mortgage payments coupled with lower home values will discourage consumer spending particularly on consumer durables. Further, the employment effects of the housing industry are considerable. In the last three years, 30% of the job growth was due directly or indirectly to housing (Roubini, 2006). The construction and real estate mortgage job loses could be considerable if housing prices drop substantially. In sum, a real estate collapse could tip the economy into a recession. We shall attempt statistically to determine whether the US housing market experienced a bubble after 2000. The detection of a bubble is important for two reasons. First, we may prepare for the effects of the bubble bursting on the remainder of the economy. Second, early detection of a bubble allows the regulators to "remove the punch bowl" and avert some of the difficulties of a housing market crash.

LITERATURE REVIEW

The widely accepted efficient market theory claims that financial asset prices reflect all the publicly available information at all times. This denies the possibility of a bubble. While some may believe prices are too high relative to fundamental factors, according to the theory they are wrong. Because investors recognize immediately if the price of anything is too high (or too low) and respond by selling (or purchasing), the asset until the over-(under) pricing is eliminated. A mountainous body of academic research supports this view, Fama (1970).

Nevertheless, the efficient market theory has been subject to much serious criticism, Shiller (2004). Furthermore, much of the research focused on financial assets. The efficient market theory assumes that investors can sell an asset short to eliminate overpricing. Real estate is a real and illiquid asset. During the period of the housing price run-up there was no mechanism known to us for shorting a residential home. A futures market for housing is a relatively recent innovation. These markets do not function well enough to fulfill the assumptions of the efficient market theory. Thus, we should not dismiss the possibility of a housing bubble out of hand.

No single explanation for real estate bubble or bubbles in general has achieved acceptance. Since this literature is too vast for this paper, we will focus mostly but not exclusively on popular explanations and indicate why they do not provide sufficient explanations for bubbles. The greater fool theory purposes that bubbles are driven by optimistic investors (fools) who buy with the expectation of selling to even more optimistic investors (bigger fools) at a profit. The bubble finally bursts when the greater fool fails to find another buyer. Thus, the greatest fool suffers a loss and the bubble unwinds. Many of the popular explanations of the housing bubble rely on the greater fool theory. Consider the following: Supposedly, the American dream of home ownership fuels a frenzy home buying unjustified by economic reality. Americans believe home prices cannot fall, despite the reality that they sometimes do. The proliferation of TV real estate reality programs and other promotions create froth in home prices.

Unfortunately, the evidence is not favorable to this popular theory. Lei et al (2001) find that bubbles occur even in the absence of overconfidence. They attribute bubbles to irrational behavior. Levine and Zajac (2007) show that even without uncertainty or the possibility of speculation bubbles appear even in well-organized experimental markets. They "find no support for the greater fool explanation." They believe the cause lies in "the institutionalization of social norms." People copy the behavior of others.

Another explanation that may be especially relevant for the real estate market is that bubbles result from increases in the money supply. When the money supply grows too rapidly the concomitant lower interest rates discourages investment in savings accounts and encourages investment in financial assets and real estate. The low interest rates reinforce this by encouraging the use of leverage. The bubble ends when

the money supply contracts. However, Smith, Suchneck, and Williams (1988) find bubbles even in experimental markets that preclude money supply changes.

METHODOLOGY

The first step in supporting the presence of a bubble is to provide a clear definition of the term. Case and Shiller (2003) define it as:

> …a situation in which excessive public expectations of future price increases cause prices to be temporarily elevated.

They argue that during a bubble homeowners spend more on a home than they otherwise might because they a believe that the house will be worth even more in the future or that the home will soon be even more unaffordable. In sum, expectations of higher housing prices make housing an irresistible investment, high return and virtually risk-free. Since housing prices cannot go up indefinitely, when they stop rising, expectations are revised and the bubble bursts.

According to Case and Shiller's (2003) definition, a bubble is a psychological construct, so it is natural to conduct investor surveys to ascertain the presence of a bubble. They surveyed recent homebuyers in three cities that apparently experienced a rapid price increases along with those in one city that has not. They find that homebuyers in 2003 expect home prices to increase in both the long and short run. Unlike their 1988 survey, this was even true for the city with the slow housing price increases. This clearly supports the presence of a bubble as defined by Case and Shiller (2003). They also found a strong belief that prices are sticky downward. Furthermore, they found buyers in the glamour cities inclined to support simplistic theories that might translate into behavior that supports a bubble. For example, buyers in the glamour cities were more likely to believe "that desirable real estate just naturally appreciates." (p. 325). This is naïve because the argument suggests only that desirable real estate will have a high price but says nothing about how that price will change. Yet the prevalence of this belief might explain how a bubble continues without the participants recognizing it. Their survey data leave the reader with the strong impression of the presence of a housing market bubble.

As Case and Shiller (2003) recognize many economists oppose drawing conclusions on economic behavior based upon consumer's thought processes. Economists prefer to focus on consumer actions. Thus, they also reinforce their case by presenting some statistical evidence. They find that regressions including income and other fundamentals explain housing prices in all but the eight glamour states. In those states, the regressions underestimate home prices during the 2000-2002 period.

In this paper, we will present additional statistical evidence in favor of the plausibility of a bubble in the real estate market. This should shed more light on the subject for those both sympathetic and unsympathetic to survey research. The cointegration methodology developed by Granger and Engle (1987) provides a method of testing for the existence of a bubble. Diba and Grossman (1988) and Campbell and Shiller (1987) apply this methodology to test for the presence of a bubble in the equity market. We use this methodology to check for a bubble in the housing market.

To explore the existence of a housing bubble we examine the stability of the underlying relationship of home prices and the economic forces that determine them. A relationship suddenly becomes unstable during a period of rising home prices is consistent with the presence of a bubble. Cointegration is well suited to test for this. Cointegration tests for the long run relationship among variables. Cointegration implies two variables share a common stochastic trend. A common stochastic trend does not simply mean that they move upward or downward together, but rather that the variables may share both prolonged upward and prolonged downward movements.

Suppose housing prices are cointegrated with an economic variable and a bubble develops in the housing market then housing prices rise without a corresponding rise in the variable. This implies the severing of a long-term relationship between housing prices and the variable. In other words, the cointegration should cease. In summary if there was a bubble in the housing market after 2000, we should be able to find variables, which were cointegrated with housing prices before 2001, and no longer remain cointegrated.

TESTING FOR COINTEGRATION

The first step in establishing cointegration is to test the variables for stationarity. Roughly speaking, a time series is stationary if the series fluctuates with a constant variation about a mean, which remains constant over time. A stationary variable has a propensity to return to and frequently cross the mean. When it is far from the mean, it is more likely to move towards rather than away from the mean. Since a stationary variable shows, a tendency to revert to a constant mean it exhibits no trend. Thus, only variables, which are non-stationary, may be cointegrated. To establish nonstationarity we employ the ADF (Augmented Dickey Fuller) test and the Phillips Peron test. The second step employs Johansen-Juselius procedure (1990) to test for cointegration. This test for cointegration entails estimating the following cointegration regression:

$$Y_T = a + bx_t + u_t \quad (1)$$

Where for our 2 variable case y_t is the housing price index and x_t represents a fundamental variable that is hypothesized to influence housing prices, a and b are regression parameters, and u_t is the error term.
We may rewrite equation (1) as:

$$\Delta y_t = a + (b-1)\, x_{t-1} + u_t \quad (2)$$

If the variables are cointegrated then b, the slope term, equals one. Thus, the test revolves around the magnitude of the estimated b coefficient. If the null hypothesis b=1 is not rejected, then equation (2) is a stable autoregressive equation and the data are inconsistent with cointegration. If we find the data consistent with b=1, then equation (2) is a random walk, in which case the variables are cointegrated.

The Johansen- Juselius procedure views this equation in a matrix format. It tests the magnitude of the b coefficient indirectly. It tests by measuring the rank of the coefficient matrix. In the two variable case we consider, the matrix is a 1x1 consisting of the b-1 term. The rank can only be zero or one. The magnitude of the eigenvalue of the matrix determines the rank of the matrix. The rank of the matrix equals the number of nonzero eigenvalues. If the eigenvalue is not statistically different from zero, then the rank is zero. Then we reject the null hypothesis (b=1); this leads to the conclusion that the variables lack cointegration. If the eigenvalue is statistically different from zero, then we conclude the rank of the matrix equals one. This implies the variables are cointegrated.

Data

Quarterly data is used and the study covers the period 1975Q1 -2007Q2. We employ the U.S. Office of Federal Housing Enterprise Oversight (OFHEO) Home Price quarterly index to measure housing prices. The index is not seasonally adjusted. A difficulty faced in the construction of a housing index is the extreme lack of uniformity of the commodity. To control for this OFHEO adopts the Case-Shiller geometric-weighted repeat sales procedure. Calhoun (1996) provides the details of the procedure. Next, we consider a series of seven variables, which reflect the fundamental economic forces determining housing prices. The most important of these is income. Case and Shiller (2003) conclude that in non-bubble markets income explains most of the rise in housing prices. We employ two separate measures of

income. The first is the mean of the middle quintile of the income distribution, denoted as the Middle Fifth. Second, we use the mean of the highest quintile of the income distribution denoted as Top Fifth. This attempts to account for the possibility that the wealthiest segment of the population influences housing prices disproportionately because of their greater mobility. The U.S. Census Bureau, *Historical Income Tables-Families* (all races), and the National Association of Realtors provided this data.

The Mortgage Rate represents a strong influence on consumer demand for housing. We obtained 30-year conventional mortgage rate (fixed rate, first mortgages) from the Board of Governors of the US Federal Reserve System. The civilian unemployment rate measures the state of the economy. The U.S. Bureau of Labor Statistics provided the seasonally adjusted percent of civilian unemployment. We converted the monthly data for both variables to quarterly data by a simple mean. The Homebuilders Stock Index provides an indication of the state of the housing market. Merrill Lynch supplied a capitalization-weighted, price level index of homebuilding stocks based on stocks included in the S&P 500 stock index.

The final variables measure the ability of consumers to handle mortgage debt. The Household Debt Ratio is the ratio of household credit market debt outstanding to annualized personal disposable income. This data also came from the Board of Governors of the US Federal Reserve System. The Housing Affordability Index for all homebuyers (HAI) measures whether or not a typical family could qualify for a mortgage loan on a typical home, assuming a 20% down payment. We define a typical home as the national median-priced, existing single–family home as calculated by NAR. In its final form used here, the HAI is essentially "median family income divided by qualifying income." The index is interpreted as follows: a value of 100 means that a family with the median family income (from the U.S. Bureau of the Census and NAR) has exactly enough income to qualify for a mortgage on a median-priced home. National Association of Realtors (NAR) provides the data. In this research, the monthly HAI values result from quarterly samples. For more details on the exact calculation, go to www.realtor.org.

PRELIMINARY ANALYSIS

Table 1 shows the simple correlation among the variables during the period before rapidly escalating home prices (1975Q1 – 2000Q4). During this period the fundamental variables exhibited a strong association with the lowest correlation coefficient of .42 between the HAI and the Mortgage Rate and the high of -.96 (both in absolute terms) between the Income of the Top Fifth and the Household Debt Ratio. More importantly, all the variables displayed strong correlation with home prices and the signs were as expected. Mortgage rates have the lowest correlation of -.48 with home prices while the Income of the Middle Fifth has the highest correlation of .88 with home prices.

Table 2 shows similar results for the whole period (1975Q1 – 2007Q3). The Homebuilder Stock Index and The Housing Affordability Index correlation coefficient of .23 was the lowest; and a correlation of -.94 between The Household Debt Ratio (Debt/Income) and Top-Fifth was the highest. The correlations between the variables and Housing prices were strong for the whole period. The Housing price index had the lowest correlation of .26 with HAI and the highest correlation of .92 with both Household Debt Ratio and Homebuilder Stock Index. Unemployment and The Housing Affordability Index were the only variables to suffer a serious decline (in absolute terms) in correlation with housing prices during the period of escalating housing prices (from -.83 to -.56 for the former and .49 to .26 for the latter). The correlation between income variables and housing prices declined minimally (less than .05). The other variables obtained stronger correlations. Based upon the correlation analysis there is little to support the hypothesis that the relationship housing price with fundamental economic variables was substantially altered during the period of housing price rises.. This occurs because correlation analysis is ill suited to detect this shift for non-stationary variables. Cointegration offers a better alternative to which we turn to in the next section.

Table 1: Correlation Matrix 1975-2000

	Mortgage Rates	Unemployment	Debt/Income	Housing Affordability Index	Home Builder Stock Index	Middle Fifth	Top Fifth	Housing Price Index
Mortgage Rates	1							
Unemployment	0.59	1						
Debt/Income	-0.644	-0.77	1					
Housing Affordability Index	-0.96	-0.6	0.71	1				
Home Builder Stock Index-Builder	-0.42	-0.48	0.71	0.46	1			
Middle Fifth	-0.64	-0.9	0.83	0.65	0.65	1		
Top Fifth	-0.66	-0.79	0.96	0.73	0.76	0.91	1	
Housing Price Index	-0.48	-0.83	0.84	0.49	0.58	0.88	0.82	1

Table 2: Correlation Matrix 1975-2005Q2

	Mortgage Rates	Unemployment	Debt/Income	Housing Affordability Index	Home Builder Stock Index	Middle Fifth	Top Fifth	Housing Price Index
Mortgage Rates	1							
Unemployment	0.65	1						
Debt/Income	-0.74	-0.69	1					
Housing Affordability Index	-0.9	-0.58	0.52	1				
Home Builder Stock Index-Builder	-0.55	-0.42	0.85	0.23	1			
Middle Fifth	-0.76	-0.84	0.91	0.59	0.71	1		
Top Fifth	-0.77	-0.78	0.94	0.66	0.7	0.95	1	
Housing Price Index	-0.59	-0.56	0.92	0.26	0.92	0.83	0.79	1

EMPIRICAL RESULTS

For all the variables, both the ADF tests and the Phillips-Perron statistic support non-stationarity at a five per cent level of significance. (The results are not shown but available upon request). We conclude that all the variables are non-stationary. Next, we examine whether home prices and the seven fundamental variables are cointegrated. We accomplish this by examining a cointegrating regression for each of the seven variables with home prices. Table 3 presents the results of these cointegration tests for the1975Q1-2000Q4 period. The Trace Statistic Test shows that for three of the seven variables, Top Fifth, Middle Fifth and the Unemployment Rate, we may reject the null hypothesis of no cointegration at a five per cent level of significance. Furthermore, we may reject the null hypothesis of no cointegration at the 10% level of statistical significance for one additional variable the Homebuilders Stock Index. Thus for the period preceding the run up in home prices there appears to have been a strong link between home prices and both the income variables and the Unemployment Rate and a marginal link with the Home Builders Stock Index.

Table 4 presents the results of these cointegration tests for the period 1975-2007Q3. The trace tests indicate the eigenvalues are not statistically distinguishable from zero in any equation at the 5%. However, the P-value for the middle fifth of income was .0502. Recognizing the belief that the bubble burst in late 2005 we did the cointegration test for the period 1975-2005Q2. The P-value (hypothesized no. of CE(s) = none) for Home prices vs. middle fifth of income was 11%. Although we did not display the results, we cannot reject the hypothesis of no cointegration for any of the other fundamental variables during this period. (The results are available upon request.) This suggests that in the post 2005 period the normal relationship between Home Prices and income was reasserting itself. This result suggests that the

linkage between Home Prices and the fundamental variables weakened substantially after 2000. The evidence is consistent with a real estate bubble.

Table 3: Cointegration Tests 1975-2000

Cointegration between	Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob. **
Home Prices vs. Household Debt Ratio	None	0.09	14.72	25.87	0.60
	At most 1	0.05	5.28	12.52	0.56
Home Prices vs. Housing Affordability Index	None	0.13	20.34	25.87	0.21
	At most 1	0.07	7.06	12.52	0.34
Home Prices vs. Mortgage Rate	None	0.10	18.16	25.87	0.33
	At most 1	0.08	7.73	12.52	0.27
Home Prices vs. Homebuilders stock index	None	0.15	25.69	25.87	0.05
	At most 1	0.09	9.49	12.52	0.15
Home Prices vs. Unemployment Rate	None *	0.15	25.90	25.87	0.05
	At most 1	0.09	9.66	12.52	0.14
Home Prices vs. mean of Middle fifth of Income	None *	0.20	32.58	25.87	0.01
	At most 1	0.10	9.90	12.52	0.13
Home Prices vs. mean of Top fifth of Income	None *	0.21	32.29	25.87	0.01
	At most 1	0.08	8.70	12.52	0.20

The table shows the eigenvalues of a regression of housing prices on the seven fundamental economic variables. For each eigenvalue we can find an associated characteristic root. The natural log of the characteristic root times the times the number of quarters (156) yields the trace statistic. This statistic tests the null hypothesis of no cointegration. We may reject this hypothesis if the p-value is less than .05 for the case of the no. of CE(s) [co integrating errors] equals none. * denotes rejection of the null hypothesis of no Cointegration at the 0.05 level
** denotes the p-value

Table 4: Cointegration Tests for the Whole Period 1975-2007Q3

Cointegration between	Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob. **
Home Prices vs. Household Debt Ratio	None	0.07	7.68	15.49	0.5
	At most 1	0	0.29	3.84	0.59
Home Prices vs. Housing Affordability Index	None	0.12	13.64	15.49	0.09
	At most 1	0.01	0.89	3.84	0.35
Home Prices vs. Mortgage Rate	None	0.1	10.63	15.49	0.24
	At most 1	0	0.42	3.84	0.52
Home Prices vs. Homebuilders stock index	None	0.1	12.28	15.49	0.14
	At most 1	0.02	1.78	3.84	0.18
Home Prices vs. Unemployment Rate	None	0.1	12.13	15.49	0.15
	At most 1	0.02	1.81	3.84	1.18
Home Prices vs. mean of Middle fifth of Income	None	0.13	15.48	15.49	0.05
	At most 1	0.02	2.02	3.84	0.16
Home Prices vs. mean of Top fifth of Income	None	0.09	8.85	15.49	0.38
	At most 1	0	0.03	3.84	0.87

The table shows the eigenvalues of a regression of housing prices on the seven fundamental economic variables. For each Eigenvalue we can find an associated characteristic root. The natural log of the characteristic root times the times the number of quarters (156) yields the trace statistic. This statistic tests the null hypothesis of no cointegration. We may reject the hypothesis if the p-value is less than .05 for the case of the no. of CE(s) [co integrating errors] equals none. Trace test indicates no Cointegration at the 0.05 level for any variable
*** denotes the p-value*

CONCLUSION

The statistical evidence we presented supports the Case and Shiller (2003) survey conclusions of the presence of a real estate bubble in the housing market. Our study uses a national housing index. Case and Shiller (2003) use local housing indices. They find the bubble limited to a few cities. They find that income has the strongest influence on housing prices. We also find a strong relationship between income and housing prices during the pre-bubble period. This relationship severed during after 2000. Case and

Shiller (2003) speculate that unemployment rates are a key precipitating factor of the housing bubble. Thus, upward (downward) movements in the unemployment rate will decelerate (accelerate) home prices. In contrast, we find a link between home prices and unemployment during the pre-bubble phase. After 2000, the linkage has disappeared. We were somewhat surprised to find the lack of a cointegration relationship between housing prices and mortgage rates.

Since we were able to find the cointegration vanished as early as 2005Q2, we believe cointegration can serve as an early bubble detection system. Alan Greenspan's (former Fed Chairman) statement, "I really didn't get it" (the housing bubble and subprime crisis) until late 2005 or 2006" supports the usefulness of this analysis Perhaps, with cointegration methodology, the Fed could have "removed the punch bowl" in a timely manner.

In the future cointegration analysis should help to identify local housing market bubbles. Since there are housing indices for local markets, this extension should prove feasible. In addition, the technique could help to verify bubbles in other asset markets.

BIBLIOGRAPHY

Calhoun, Charles, 1996, OFHEO House Price Indexes: HPI Technical Description, Office of Federal Housing Enterprise Oversight, Washington D.C.

Campbell, John and Robert Shiller, 1987, "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95 (October), p. 1062-88.

Case, Karl and Robert Shiller, 2003, Is there a Bubble in the Housing Market," Brooking Papers on Economic Activities, 2, p. 299-362.

Diba, Behzad and Herschel Grossman, 1988, "Explosive Rational Bubbles in Stock Prices," American Economic Review, 78, p. 520-630.

Dickey, David and Wayne Fuller, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," Econometrica, 49, (July), p. 1057-72.

Granger, Clive and Engle, Robert, 1987, "Cointergration and the Error-Correction Representation: Estimation and Testing," Econometrica, 55, p. 251-76.

Engle, Robert and B. Yoo, 1987, "Forecasting and Testing in Co-integrated Systems," *Journal of Econometrics*, 35, p. 143-159.

Fama, Eugene, 1970, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, p. 383-417.

Greenspan, Alan, Interview with Jim Lehrer (http: "youtube.com/watch?v=yu4wjhijomy>)". The news show with Jim Lehrer, 18 September 2007.
Johansen, Soren and Katrina Juseluis, 1990, "Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money," Oxford Bulletin of Economics and Statistics, 52, 169-209.

Lei, Vivian, Charles Noussair, and Charles Plott, 2001, "Nonspeculatative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality", *Econometrica*, 69, p. 831-859.

Levine, Sheen, and Edward Zajac, 2007, "The Institutionalization of Price Bubbles," Available at SSRN http://ssrn.com/abstract=960178.

Lowenstein, Roger, 2007, "On the Home Front; Pop Psychology," New York Times Magazine, March 18, 2007.

Phillips, Peter and Pierre Perron, 1988, "Testing for a Unit Root in Time Series Regression," Biometrika, 75 (June), p. 335-46.

Roubini, Nouriel, 2006, "The Biggest Slump in the Last 40 Years…or 53 Years," Nouriel Roubini's Blog, August 23, 2006.

Shiller, Robert, 2004, "From Efficient Markets to Behavioral Finance," *Journal of Economic Perspectives*, 17 (1), p. 83-104

Smith, Vernon, Gerry Suchenek, and Arlington Williams, 1988, "Bubbles, Crashes, and Endogenous Expectation in Experimental Spot Asset Markets," *Econometrica,* 69, p. 831-59.