EKONOMIHÖGSKOLAN
NATIONALEKONOMISKA INSTITUTIONEN
LUNDS UNIVERSITET

KANDIDATUPPSATS
AUGUST 2011

THE IMPACT OF VACANCY RATES ON HOUSE PRICES – AN ECONOMETRICAL ANALYSIS

AUTHOR

MÅRTEN BERG

SUPERVISOR

FREDRIK NG ANDERSSON

Executive Summary

Title The impact of vacancy rates on house prices – An econometrical analysis

Author Mårten Berg

Supervisor Fredrik NG Andersson

Level Bachelor

Abstract The purpose is to examine the relationship between the rates of vacant houses with the price of housing in different regions within the United States; the Northeast region, the Midwest region, the South region and the West region. Results are analyzed through economic theory on vacancy and house prices developed by (Wheaton, 1990). The model proposed examines annual GDP, Interest-rates and vacancy rates on the price of housing from 1963 to 2010. The data is region specific and collected from the US census, the Bureau of economic analysis and the Federal Reserve. The method used is the Engle-Granger test for cointegration. The results are mixed and while the Northeast exhibit most coherence with theory it is not in the normal inverse relationship with prices. Instead the prices increase as the rate of vacancy increases. In essence, the case for including vacancy rates in more models of house prices is relatively weak and variables such as real income have shown better cointegration results.

1. Introduction

The recession that began in 2007 (The National Bureau of Economic Research, 2010) is widely attributed to the 2007 financial crisis when banks and investment banks all over the world found themselves insolvent. A great deal of financial innovation, in years prior, made the markets for financial instruments with home mortgages as underlying asset a major revenue source for financial institutions. But the revenue of the instruments depended on the home owners' ability to pay their mortgage and if they did not, that house prices would increase so the property could be resold at a higher price. Only if housing prices declined significantly across the board would the financial instruments lose value. Such a large nationwide decline in housing prices had never occurred in the United States from the beginning of housing price data collection. Not until the 2007 financial crisis.

Recessions after real estate bubbles is not a new phenomenon (Makin, 2006), one of the most notable being the 1989 crash and culmination of the bubble in Japan. Since the bubble burst, Japan has experienced at least a lost decade, possibly two, when GDP didn't grow at all. The implications of a major drop in asset prices on a national level can be dire. Explained by (Koo, 2009), a decline in asset prices makes the balance sheet of firms and citizens unbalanced which leads to a shift from profit maximizing to debt repayment. Further, in our current monetary system a net repayment of debt decreases the amount of money in circulation, causing the economy to contract. Efficient economic policy, fiscal, monetary and structural, could help alleviate the strain on the economy suffering from a general decline in asset prices. But the causes of asset bubbles are perhaps an even more complex area in economics.

The driving factors for house prices have been researched in previous studies and a number of variables have been tested through econometrical analysis. Real income and declines in

interest rates have been shown to possess a positive effect on house prices by (McCarthy & Peach, 2004) (Case & Shiller, 2003). The financial structure also impacts as shown by (Tsatsaronis & Zhu, 2004), they found availability of bank credit and the prevalence of floating-rate mortgages to have a significant effect. Holly *et al* (Holly, Pesaran, & Yamagata, 2010) added population growth as a significant positive effect on house prices. These factors where further confirmed by Cameron *et al* (Cameron, Muellbauer, & Murphy, 2006) who found population growth, real income growth, low housing construction and generally low interest rates on mortgages to have driven house prices in the UK. If a variable is a fundamental driver there is long-term cointregration between the variable and house prices. However, the existence of cointegration between income and house prices is not clear. Both (Meen, 2002) and (Gallin, 2006) looked for the same cointegration and did not find significant results.

The amount of economic data available for study today is greater than ever, which gives the option of examining previously overlooked trends in the years leading up to asset bubbles. One statistic mentioned in the large debate over how the crisis in 2007 could have been better predicted is the vacancy rate of homes. Home vacancies in the United States increased dramatically in the years preceding it. And the US census bureau has since 1963 collected the vacancy rate in the United States as a whole as well as in the regions: Northeast region, Midwest region, Southern region and the West region. Economic theory regarding vacancy, (Wheaton, 1990) suggest a strong inverse relationship between vacancy rates and housing prices under normal circumstances. Further, the theory explains that a positive effect on housing prices is possible if the vacancy rate increases from a very low percentage, thereby unlocking the market for houses. More people would then be interested in switching homes and their entrance into the market is the force that would drive up prices.

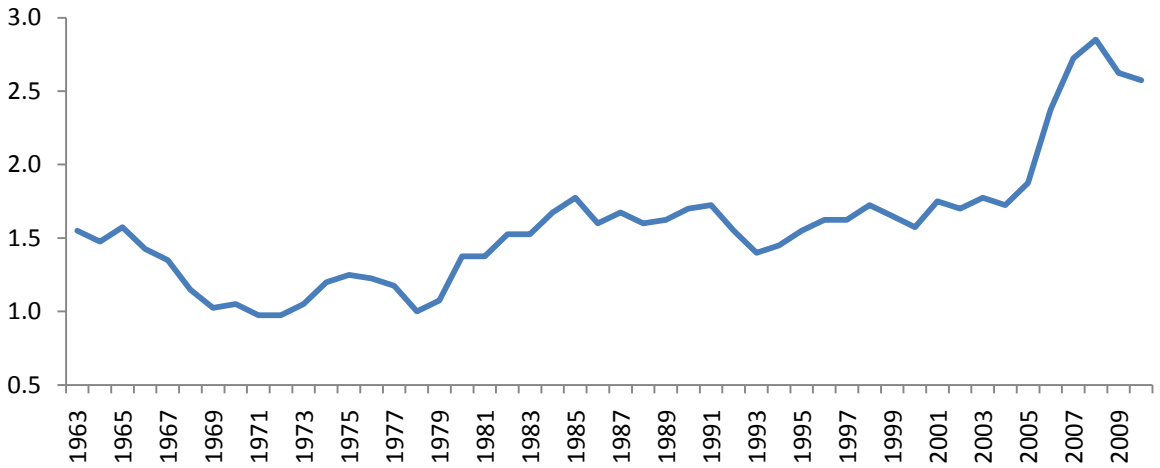

Figure 1: US Home vacancy rate, Source: US census bureau

In this paper I will look for cointegration between house prices and vacancy rates using an econometric cointegration test developed by (Engle & Granger, 1987). Testing is done on the four regions of the United States as defined by the US Census bureau: the Midwest region, the Northeast region, the South region and the West region. The data spans over 48 years from 1963 to 2010. The fundamentals of all prices lay in basic supply and demand relationship and the amount of unsold houses can be seen as an indicator of that relationship. If house prices and vacancy rates are cointegrated it means that there is a long-term relationship between the two. The variables may drift apart for some time but in the long run they return to each other.

Modeling of regional house prices has been in numerous ways. One of them is a general model proposed by a Dutch analysis (Van Dijk, Franses, Paap, & Van Dijk, 2007). Their model uses a latent time-series approach. Impulse-response functions are used to test the impact of potential changes to real GDP and the contagious effect regarding regions close to each other. Even here vacancy rates were not included in the model but results show strong evidence for regional differences with specific price dynamics. Another way to examine housing prices was presented by Cameron *et al* (Cameron, Muellbauer, & Murphy, 2006)

using inverted demand functions. They look at a number of different variables to explain the British house market from 1972-2003 when house prices rose strongly. Holly *et al* (Holly, Pesaran, & Yamagata, 2010) constructed a fully heterogeneous model where different regions have different parameters.

As with the other studies, this study finds mixed results on cointegration for fundamentals in determining house prices. Vacancy rates exhibits weak results for cointegration with house prices across the regions. Northeast exhibits a cointegrating relationship; South and West present mixed results from the cointegration test and the Midwest show no cointegration relationship.

The remainder of this paper is structured with Section 2 as a brief overview of the US housing market with relevant historical trends. Section 3 then explains the model for the cointegration tests and the data used. Section 4 provides the data used in the regression analysis as well as result of the unit root tests. Section 5 shows the results and section 6 concludes with onward discussion. Following is a bibliography and an appendix with the full regression results and the data used.

2. US housing market

Vacancy rate theory for homes, developed by (Wheaton, 1990), characterize the housing market by noting that real estate markets have a structural vacancy, similar to structural unemployment in labor models. This structural vacancy differs from area to area and depends on variables such as new construction and the activity level in the area itself. Market conditions in the single family home market are also characterized by switching on the part of

current homeowners. A large part of homebuyers are also sellers in the same market which means that the real estate market can be described in terms of matching. When homebuyers look for a new home there is a search period that is costly both in time and lost income due to the length of the search required.

But there are differences between the labor market and the real estate market. Structural unemployment, as described by (Lucas & Prescott, 1974), features workers moving from job to job leaving an employment opportunity open before they find new employment. The housing market meanwhile features homebuyers who usually keep two houses for a short period of time before they can sell their old home at an acceptable price. This period of time that homebuyers need to pay for two homes is an additional cost that is linked with the search time of the people who will buy their old home. The market search time in turn is dependent on the vacancy rate.

The conclusions of the vacancy rate theory by (Wheaton, 1990); suggest that even relatively minor changes in supply or demand with their impact on vacancy can have a significant effect on house prices. For market participants already within the housing market looking to switch homes dramatically increased prices only increases their transaction costs. Overall, the relationship between vacancy and house prices is found to be negative. However the existence of structural vacancy means that a certain level of vacancy is needed for an efficient market. When the vacancy rate is below the structural vacancy the market slows down and participants cease efforts to look for new homes. Thus an increase in vacancy from a very low percentage has the possibility of unlocking the market. Search costs will then become lower and activity in the market will increase leading to an increase in prices.

One of the most widely used metric for house prices is the Standard and Poor's/Case-Shiller index. The index developed by Case & Shiller uses a "repeat sales" method of measuring house prices which uses sales information on houses that have been sold at least twice in a four quarter period. This provides additional assurance of the accuracy of the valuation (Standard & Poor, 2011). Following is the annualized S&P Case-Shiller index for the whole US market from 1987 to 2010 with the first quarter of 2000 being 100:

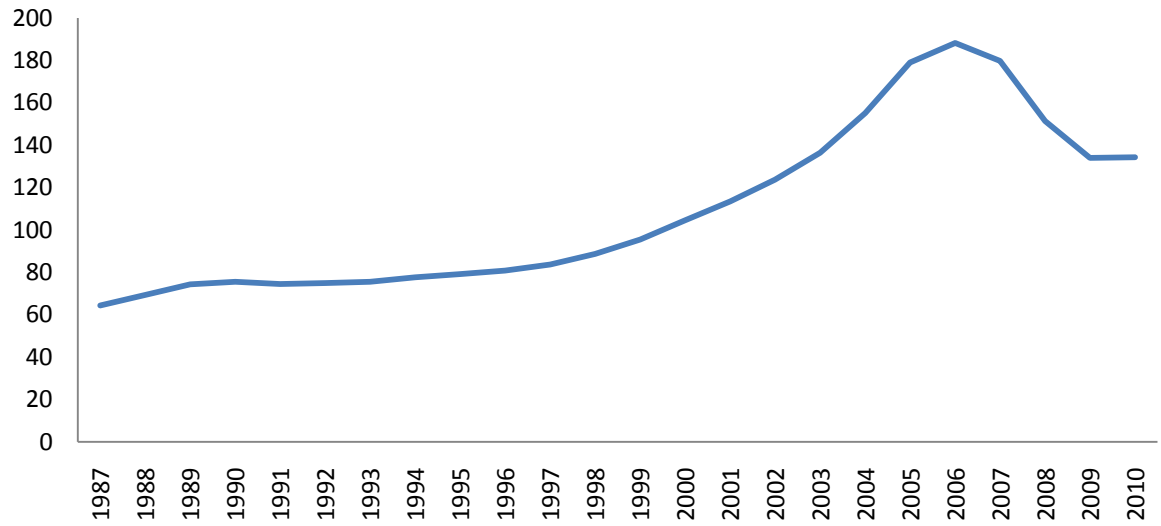

Figure 2: Annualized S&P Case-Shiller Index for the whole US market (2000=100), Source: Standard and Poor

The federal government encourages homeownership through a number of enterprises and policies. Government sponsored enterprises like Fannie Mae (Federal National Mortgage Association) and Freddie Mac (Federal Home Loan Mortgage Corporation) among other things purchase mortgages and turns them into securities, then sell them on the secondary market. Their purpose is to improve liquidity and keep rates on mortgages low in the economy. Regulations are in place to ensure that the enterprises only acquire loans of sufficient standards which include income-to-debt ratios and proper documentation. Before 1981 second mortgages were not allowed by these regulations and in 2000, worries about the impact of subprime mortgages added restrictions on Freddie Mac and Fannie Mae on how

risky their purchases were allowed to become (Koppell, 2001). However, these restrictions were removed in 2004 (Wallison & Calomiris, 2009).

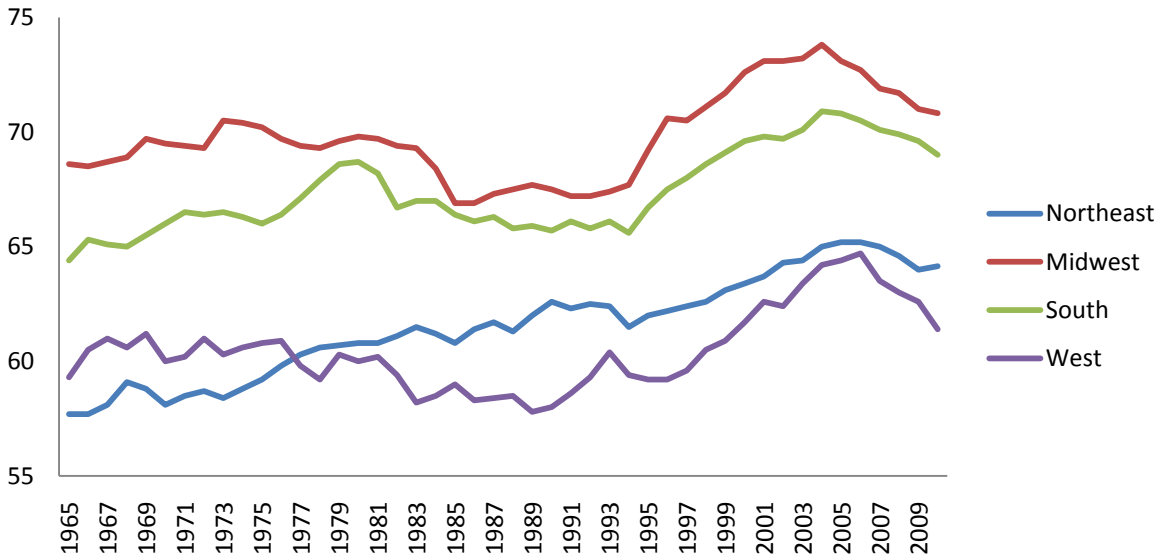

Figure 3: US homeownership rate by census defined region, Source: US census bureau

Other significant policies by the federal government include a mortgage interest tax deduction on the condition that it is the primary residence of the borrower and the Community Reinvestment Act. The Community Reinvestment Act of 1977 has since encouraged banks and other lending institutions to increase the available amount of credit in the areas around the bank's branch offices. The idea of the legislation is to counter ethnic and lower-income discrimination when reviewing applicant borrowers. This kind of discrimination was seen as a major contributor to observed urban decay, even in wealthy cities, and the Community Reinvestment Act was conceived under pressure from community interest groups. But the impact of the Community Reinvestment Act on house prices is probably marginal at best due to the low number of loans made under its specifications (Raphael & Stuart, 2006).

The Savings and Loans crisis in the late 1980s and 1990s was the result of a major collapse among financial institutions in the United States. One of the contributing factors was the removal of tax benefits for housing projects, effectively ending construction as a safe haven, in 1986. This resulted in reduced construction for many years to follow but at the same time substantial drop in house prices. The general crisis in the financial institutions combined with a radical shift in the interest rate policy at the Federal Reserve caused interest rates to rise. House prices remained substantially under the trend for a long time (Diamond & Lea, 1992).

3. Model

There are a few recognized models for researching regional house prices. One presented by (Abraham & Hendershott, 1996) is structured as an identifier of housing bubbles with coefficients that behaves differently compared to the difference between equilibrium price and the actual price. The model also includes construction costs in addition to standard variables such as interest rates and income. Another model presented by (Malpezzi, 1999), measures house prices directly to the real income ratio. It further includes variables such as population, interest rates and a measure of housing regulation stringency. Both these models operate under an error-correction assumption which in turn relies on a underlying cointegration between house prices and fundamental variables. Research, such as (Gallin, 2006), suggest that this assumption may not be true and even if it was true cointegration between house prices, income and population cannot be verified. Without verification, forecasts derived from error correction models are reasonably therefore suspect.

Instead, I propose a simpler model for testing cointegration for vacancy rates inspired by (Meen, 1999). This model uses GDP as an aggregate proxy for population and income.

Added interest rates and an intercept dummy if needed to correct for problems in the US credit market, the model does provide an indication of the value of vacancy rates as a macro variable. The model for the estimation is therefore constructed as follows:

$$p_{it} = \alpha_i + \beta_{i1} y_{it} + \beta_{i2} I_{it} + \beta_{i3} v_{it} + Dum_{it} + u_{it} \quad (1)$$

Where p_{it} = the logarithm of house prices in the i th region in year t

y_{it} = logarithm of GDP

α_i = constant

I_{it} = interest rate

v_{it} = vacancy rate

Dum_{it} = intercept dummy

Testing for cointegration is performed utilizing the Engle-Granger cointegration test (Engle & Granger, 1987). The test is performed in two steps, first a main regression is estimated through Ordinary Least Squares (OLS) and if the variables are non-stationary a unit-root test is performed on the residuals. If the residuals are found to be stationary there is a long term cointegrating relationship between the non-stationary variables. Stationary residuals indicate that the underlying trend of the non-stationary variables is the same. An Error Correction Model (ECM) is then advised if cointegration is found in order to determine the speed of adjustment of the variables back to the underlying trend.

Another commonly used test is the Johansen test (Johansen, 1991) for cointegration. The Johansen test uses a maximum likelihood estimation method which tests for the number of cointegrating relationship in the model. The Johansen approach allows for multiple cointegrating relationships in the model and may therefore achieve clearer results if those are

present. However, the Johansen test requires a large sample and if there is only one cointregrating relationship it is the same as the Engle-Granger procedure. Due to the limited amount of variables included in this model and the lack of a large sample of data the Johansen test will not be used. The test has also received critique that it often finds cointegration where none exists (Wickens, 1996).

4. Data

The data for house prices and vacancy rates is collected from the US census bureau, the interest rate is approximated from the 10-year Treasury bill from the Federal Reserve and GDP statistics is gathered from the Bureau of Economic Analysis. GDP is state-level data taken from BEAs estimations and aggregated to the appropriate regions defined by the US census bureau: the Midwestern region, the Northeastern region, the Southern region and the Western region. Data length is from 1963 to 2010 and includes 48 observations. When only quarterly data is available, the data is transformed to annual simple averages of the four quarters. Such is the case with the vacancy rate statistics.

The quality of Census data and the GDP estimates from the BEA compared to other statistics provided by the Case-Shiller index or the Office of Federal Enterprise Oversight on house prices is not flawless. But in order to analyze long term cointegrating relationships data length is of high importance. As an example, the Case-Shiller index only covers years starting from 1987 and onward. US census data covers years from 1963 which allow for better testing of long term trends, and the US census data is commonly used when superior sources are lacking.

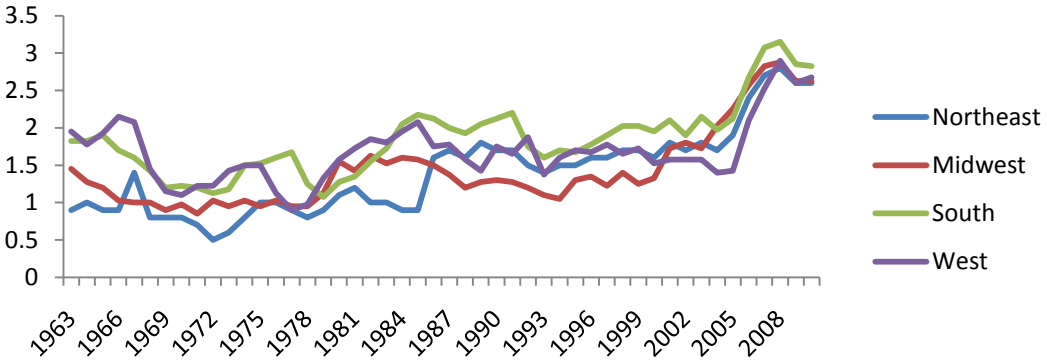

Figure 4: Vacancy Rates

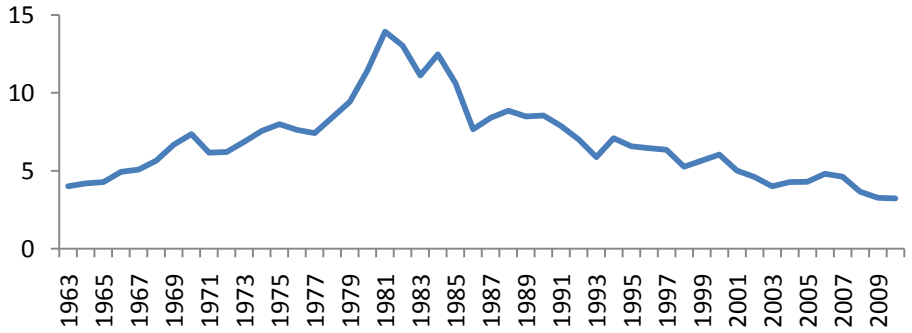

Figure 5: Interest Rate (10-year T-Bill rate)

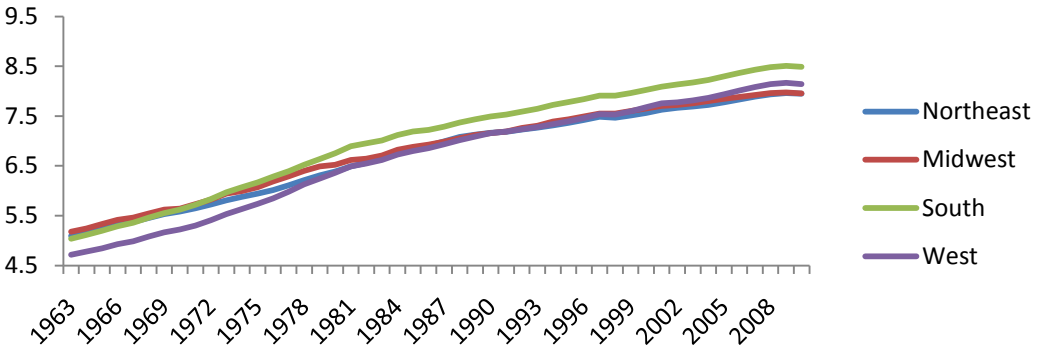

Figure 6: Log GDP

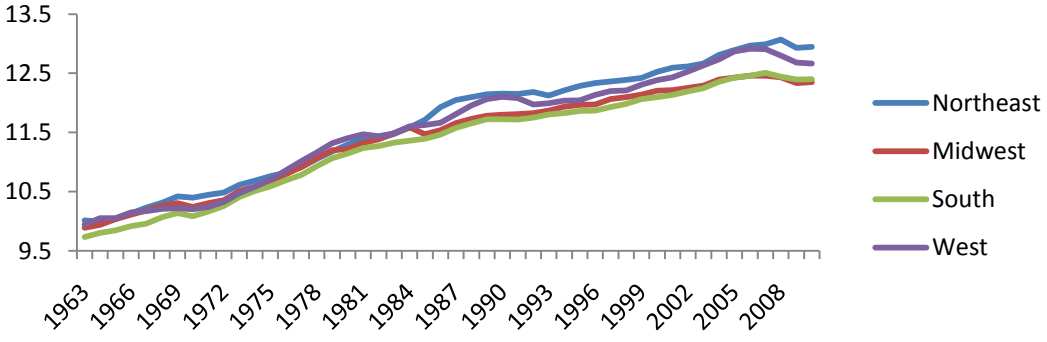

Figure 7: Log House Prices

Testing for cointegration depends on the variables being stationary or non-stationary. Stationary variables have a stable average and variance over time. In the long run the variable will always return to its average but may fluctuate in the short run. Non-stationary variables have no stable average and over time keep moving. If two or more non-stationary variables are included in a normal regression the results may show a very strong linear connection between them were none exists in reality. However, when testing for cointegration non-stationary variables are usable (Westerlund, 2005). To examine whether the variables are stationary or not a unit-root test is needed.

Table 1: Augmented Dickey-Fuller Results

Variables	T-Statistic	Prob
Log House Prices		
Northeast	-3.508508	0.986445
Midwest	-3.508508	0.997911
South	-3.510740	0.974641
West	-3.510740	0.730304
Log GDP		
Northeast	-3.508508	0.999529
Midwest	-3.508508	0.999805
South	-3.508508	0.999549
West	-3.508508	0.999782
Interest Rate		
10-year Treasury bill	-2.925169	0.676665
Vacancy Rates		
Northeast	-3.510740	0.217702
Midwest	-3.508508	0.732198
South	-3.510740	0.123743
West	-3.508508	0.556333

P values are one-sided. Log House Prices and log GDP is tested on level with intercept and a trend. The interest rate is tested on level with an intercept. Vacancy rates are tested on lag length chosen by Schwarz information criterion with trend and intercept.

To test for unit root the ADF (Augmented Dickey-Fuller) test is utilized. The test looks for autocorrelation by adding lagged values of the same series. The lagged values try to catch all autocorrelation by testing if the added lag coefficients δ are significantly different from zero. If they are, the variable is stationary. Log House Prices and log GDP is tested on level with intercept and a trend. The interest rate is tested on level with an intercept. Vacancy rates are tested with appropriate lag length chosen by Schwarz information criterion with trend and intercept.

The Augumented Dickey-Fuller test uses the t-statistic but it no longer has the ordinary t-distribution. Instead it has a special Dickey-Fuller distribution. The critical value at 5 % significance is -1.65 for a variable with intercept and trend. With only intercept the critical value is -2.86 at 5 % significance. The variables are non-stationary at 5 % significance as indicated by the P values. When using the ADF test on the residuals to test for cointegration the same Dickey-Fuller distribution cannot be used. Since the residuals are not from the original data but instead estimated by the underlying regression the critical values are acquired from the econometricians Russel Davidsson and James McKinnon, presented by (Westerlund, 2005).

5. Results

As with previous studies examining the fundamentals and their impact on house prices the results are mixed. While the Northeast seems in line with theory, although not the common inverse relationship, in the Midwest no cointegration was found. For the West and the South the results are unclear due to high auto-correlation in their respective estimations. In the Northeast region, in order to compensate for auto-correlation an intercept dummy was needed to be inserted in the underlying equation for the years 1988 to 2002. Then the following parameter results present themselves:

Table 2: Parameter results, Northeast region

Variable	Coefficient	t-Statistic	Prob.
Constant	4.7024	40.7007	0.0000
Log GDP	0.959623	39.4248	0.0000
Interest Rate	0.003033	0.6279	0.5334
Vacancy Rate	0.385057	10.9639	0.0000
Dummy	-0.110388	-4.0968	0.0002
R-squared	0.995395		
Durbin-Watson stat	1.77914		

The R-squared is 0.995395 which is very high and could indicate a spurious regression but the Durbin-Watson value is 1.77914 which is above the critical value of 1.670 for 50 observations and 4 variables. This indicates that there is none to very low auto-correlation in the model. For there to be cointegration the residuals of the regression needs to be stationary. The same ADF (Augmented Dickey-Fuller) test that was used to test the variables can be used for the residuals:

Table 3: Residual Unit-Root test, Northeast region

ADF test statistic	t-Statistic	Prob.
Test critical value:	-6.163651	0.0000
Critical value at 5% level	-4.42	

As the ADF test value is lower than the critical value at 5 % significance, cointegration can be concluded. The dummy added stretching from 1988 to 2002 can be justified due to the Savings and Loans crisis in the late 1980s. The crisis made interest rates go much higher than normal and through the restricting of credit house prices did not work as under normal conditions. Therefore it's fair to say that house prices were undervalued in the years following until the lending institutions had recovered.

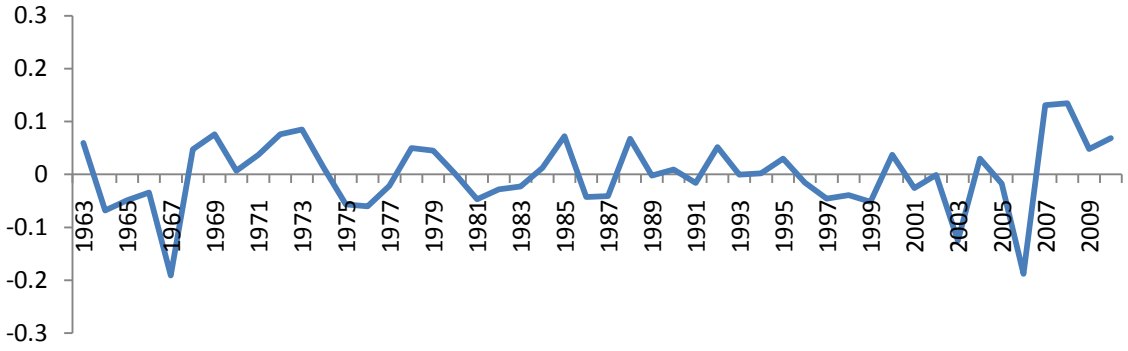

Figure 8: Residuals for Northeast with added dummy

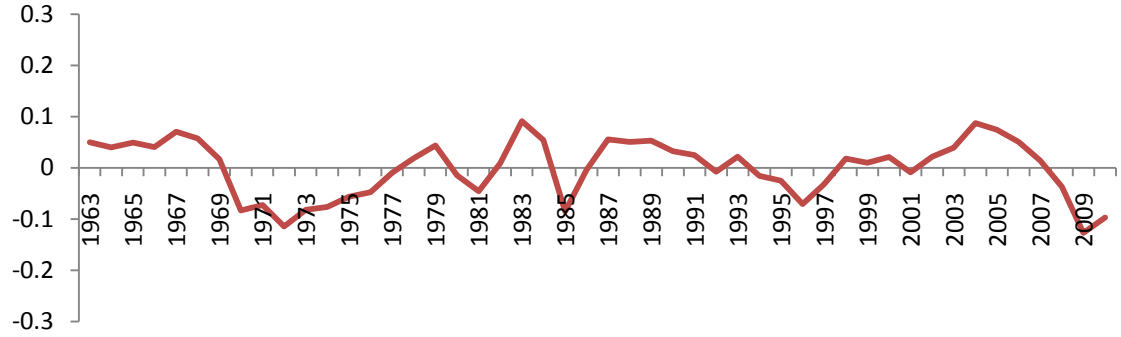

Figure 9: Residuals for Midwest

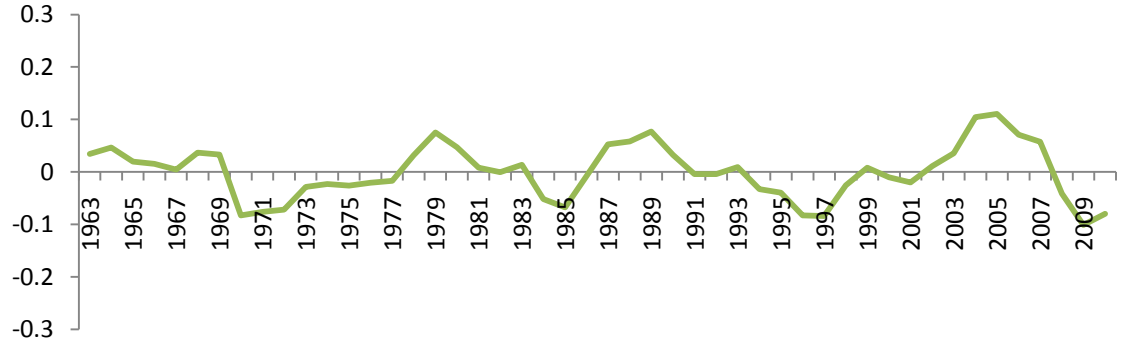

Figure 10: Residuals for South

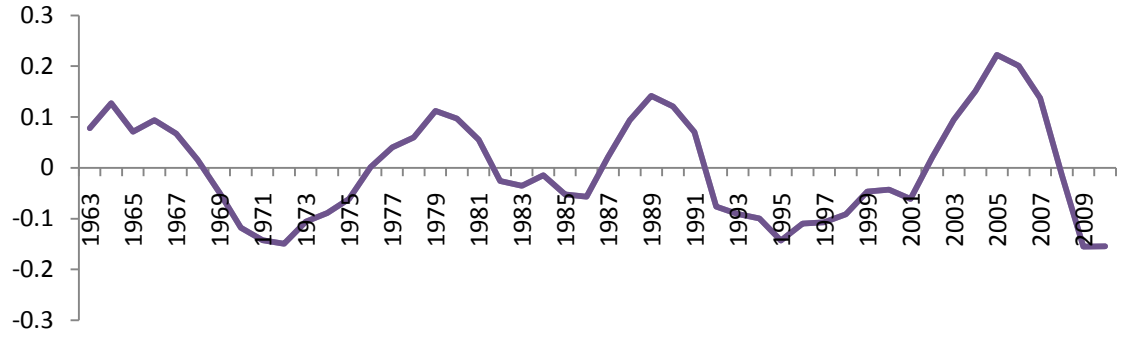

Figure 11: Residuals for West

With a shown cointegrating relationship in the Northeast region the coefficients can be viewed as the long run parameters. The coefficient for the Vacancy rate shows 0.385057 which is a positive value. This can seem counter-intuitively as the intuitive relationship should be that more vacant homes would lower house prices. Compared with theory (Wheaton, 1990) the Northeast exhibits a relationship where an increase in vacancy provides an increase in prices. This is possible under the condition that vacancy rates start so low that the housing market is frozen because the amount of available homes to move into is too low. Then an increase in vacancy unfreezes the market and as more people enter, prices rise.

The Midwest region exhibit a similar R-squared result to the Northeast region with an R-squared of 0.995391. Such a high R-squared is suspicious but not uncommon when dealing with non-stationary variables. However, the Durbin-Watson auto-correlation statistic is 0.632958 which is far below the need 1.670 and indicates a lot of auto-correlation in the model. The ADF test for unit root on the residual give the following result:

Table 4: Residual Unit-Root test, Midwest region

ADF test statistic	**t-Statistic**	**Prob.**
Test critical value:	-3.387846	0.0165
Critical value at 5% level	-4.1	

The ADF test result shows a t-statistic of -3.387846 which is higher than the needed critical value of -4.1 on 5 % significance level. Therefore no cointegrating relationship can be assumed. Adding of a dummy similar to the one added to the Northeast region produces an R-squared of 0.995465 and a Durbin-Watson statistic of 0.679328. The low Durbin-Watson still point towards that much auto-correlation is still present. An ADF test on the residuals of the Midwest region estimation with added dummy gets a t-statistic of -3.58491 with a p value of 0.0099. The t-statistic is still too low and no cointegration can be proved.

For the South region, the long term regression has an R-squared of 0.996459 and a Durbin-Watson statistic of 0.561905. Low Durbin-Watson statistics indicate a significant amount of auto-correlation. The residual unit root test yields the following result:

Table 5: Residual Unit-Root test, South region

ADF test statistic	t-Statistic	Prob.
Test critical value:	-4.069464	0.0026
Critical value at 5% level	-4.1	
Critical value at 10% level	-3.81	

The ADF critical value of -4.069464 is just above the 5 % critical value of -4.1 but still below the 10 % critical value of -3.81. This indicates some evidence of cointegration albeit less significant than in the Northeast region. The present amount of auto-correlation also presents a problem for the interpretation of the results. An added dummy variable may alleviate some of the problem with auto-correlation and yield better results on the residual unit-root test. The parameter results without dummy are as follows:

Table 6: Parameter results, South region

Variable	Coefficient	t-Statistic	Prob.
Constant	5.590903	94.48186	0.0000
Log GDP	0.801934	79.42525	0.0000
Interest Rate	0.005188	1.513211	0.1374
Vacancy Rate	0.023211	0.9569	0.3438
R-squared	0.996459		
Durbin-Watson stat	0.561905		

Both an intercept dummy from 1991 to 2001 and the dummy used on the Northeast region from 1988 to 2002 failed to substantially decrease the auto-correlation. The resulting Durbin-

Watson statistics is 0.573965 for the former and 0.554630 for the latter. Even the unit-root test results remain largely unchanged with -4.083432 and -4.071239 respectively. The auto-correlation and the failure to reach the critical 5 % value add uncertainty to the results and its parameters.

The West region shows a high R-squared of 0.988417 and a very low Durbin-Watson of 0.315946. A Durbin-Watson statistic of at least 1.670 is required to assume no auto-correlation and that is far away. The result of the residual test is as follows:

Table 7: Residual Unit-Root test, West region

ADF test statistic	**t-Statistic**	**Prob.**
Test critical value:	-4.243814	0.0016
Critical value at 5% level	-4.1	

From the residual t-statistic of -4.243814 which is lower than critical value, there are indications of cointegration. But, the auto-correlation remains a problem. The parameter results are as follows:

Table 8: Parameter results, West region

Variable	**Coefficient**	**t-Statistic**	**Prob.**
Constant	5.801336	49.31685	0.0000
Log GDP	0.862707	56.91644	0.0000
Interest Rate	0.001962	0.318746	0.7514
Vacancy Rate	-0.006157	-0.152818	0.8792
R-squared	0.988417		
Durbin-Watson stat	0.315946		

A dummy was tried in order to reduce the auto-correlation. With the dummy used on the Northeast region the estimation produced an R-squared of 0.989005 and a Durbin-Watson statistic of 0.341135. Hardly any change at all. The residual test got very similar results to the estimation without the dummy variable.

To put these results in perspective, Holly *et al* (Holly, Pesaran, & Yamagata, 2010) found cointregration for real income and house prices in general when they looked at 49 states in the US over 29 years. In addition, they discovered a positive impact from population increases in the states examined. This shows that the fundamentals can reveal themselves if tested adequately and over a large enough span of regions and years. However, (Meen, 2002) and (Gallin, 2006) could not find cointegration between income and house prices when they examined statistics on a national level. A possible explanation is that they looked at a too short time span of data. But nevertheless, coherent results on cointegration for house prices with fundamentals remain elusive.

6. Discussion

Further testing for cointegration is possible through the Johansen (Johansen, 1991) cointegration test. The Johansen method uses a different method from the Engle-Granger (Engle & Granger, 1987) methodology but lacks as an intuitively appealing test procedure. It remains a very commonly used test for cointegration despite some criticism of its vector autoregression approach and therefore there is a case for including it in this analysis. Due to its complicated method and the added difficulty in interpreting the result the Johansen cointegration test was left out of this study. Another aspect to further expand upon is an Error Correction Model which test the speed of adjustment of the cointegrated variables. Proposed in the Engle-Granger two-step procedure to examine cointegration to accompany the long

term regression, it would be a good addition to this paper. However, to include it is not part of the general purpose of this paper and has therefore been left out.

One problem with identifying the relevant variables for house price models is data quality and the relative short time span of much time-series panel data. US census data is widely used in econometrical studies but often passed over when better sources are available. Census data is collected using non-optimal methods and newer, better performing indices do not have enough historical data available to base research on longer term trends on. In addition, the difference between annual, quarterly and seasonal impacts and adjustments contribute to making the data and thereby analysis less reliable. Simple annual averages have been used to add up quarterly data in a multitude of studies but that approach might hide trends. Luckily, the problem of data reliability will work on solving itself as time passes and new indices gain years but for the time being it aggravates cointegration analysis.

Most modelling on house prices follow an Error Correction framework, for example the model suggested by (Malpezzi, 1999). The methodology he presents assumes cointegration for house prices and the fundamental variables included in the model. Often used variables are income and population but others have been used as well, such as a construction cost index and a measure of how strict housing regulations are present in the area examined. But some studies that have strictly tested for cointegration without an Error Correction framework have not been able to confirm that underlying cointegration. As a result (Gallin, 2006) suggests that the Error Correction framework of most models on house prices might not be the right methodology. He argues that more testing is needed to confirm that underlying cointegration relationship. Vacancy rates can fairly be described as a fundamental variable and as the results of this study is mixed it adds to the need for further testing.

The mixed result from this cointegration analysis on vacancy rates makes it unclear if it would be prudent to include vacancy rates in new house price models. Relevant variables to include in general models should have proven cointegration or very strong theoretical backing. The theoretical backing for vacancy rate relationship with house prices is relatively mild given that the theory assumes existing structural vacancy. Only large deviations from the structural vacancy in the area will significantly affect prices. On the other hand, there is the fact that many buyers in the market are at the same time sellers. If the buyer is moving to a house of the same standard in the same area as the one he's selling, the price of the new house loses importance. This means that increase in house prices would do little to dampen demand. That is an effect overlooked in models of house prices and their search for housing bubbles.

If the fundamentals that drives house prices can be better identified, so can the situations when prices deviate from their underlying factors. Such a deviation could be an indicator of a bubble in the housing market.

Bibliography

Abraham, J., & Hendershott, P. (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research* , 191-207.

Cameron, G., Muellbauer, J., & Murphy, A. (2006). *Was There A British House Price Bubble? Evidence from a Regional Panel.* Oxford: Department of Economics, University of Oxford.

Case, K. E., & Shiller, R. J. (2003). Is There a Bubble in the Housing Market? *Brookings Papers on Economic Activity* , 299-342.

Diamond, D., & Lea, M. (1992). Housing Finance in Developed Countries: An International Comparison of Efficiency. *Journal of Housing Research* , 145-170.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Economoetrica* , 251-276.

Gallin, J. (2006). The long run relationship between house prices and income: evidence from local housing markets. *Real estate economics* , 417-438.

Holly, S., Pesaran, M., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics* , 160-173.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector. *Econometrica* , 1551-1580.

Koo, R. (2009). *The holy grail of macroeconomics: Lessons from Japan's great recession.* Singapore: John Wiley & Sons.

Koppell, J. (2001). Hybrid Organizations and the Alignment of Interests: The Case of Fannie Mae and Freddie Mac. *Public Administration Review* , 468-482.

Lucas, R., & Prescott, E. (1974). Equilibium Search and Unemployment. *Journal of Economic Theory* , 188-209.

Makin, J. (2006). *Housing and American recessions.* Washington D.C: American Enterprise Institute.

Malpezzi, J. (1999). A simple error correction model of house prices. *Journal of Housing Economics* , 27-62.

McCarthy, J., & Peach, R. (2004). *Are home prices the next "bubble"?* New York: Economic Policy Review 10.

Meen, G. (1999). Regional House Prices and the Ripple Effect: A New Interpretation. *Journal of Housing Studies* , 733-753.

Meen, G. (2002). The Time-Series Behavior of House Prices: A Transatlantic Divide. *Journal of Housing Economics* , 1-23.

Raphael, B., & Stuart, G. (2006). Do the GSEs matter to low-income housing markets? An assessment of the effects of the GSE loan purchase goals on California housing outcomes. *Journal of Urban Economics* , 458-475.

Standard & Poor. (2011). *Indices: S&P .* Retrieved 14 July, 2011, from S&P Web site: http://www.standardandpoors.com

The National Bureau of Economic Research. (20 September, 2010). *Cycles: NBER.* Retrieved 24 July, 2011, from NBER: http://www.nber.org/cycles.html

Tsatsaronis, K., & Zhu, H. (2004). What Drives Housing Price Dynamics: Cross-Country Evidence. *BIS Quarterly Review* , 65-78.

Van Dijk, B., Franses, P. H., Paap, R., & Van Dijk, D. (2007). *Modeling Regional House Prices.* Rotterdam: Econometric Institute, Erasmus University Rotterdam.

Wallison, J., & Calomiris, C. (2009). The Last Trillion-Dollar Commitment: The Destruction of Fannie Mae and Freddie Mac. *Journal of Structured Finance* , 71-80.

Westerlund, J. (2005). *Introduktion till ekonometri.* Lund: Studentlitteratur.

Wheaton, W. C. (1990). Vacancy, Search, and Prices in a Housing Market Matching Model. *The Journal of Political Economy* , 1270-1292.

Wickens, M. (1996). Interpreting cointegrating vectors and common stochastic trends. *Journal of Econometrics* , 255-271.

Appendix 1 – Full results

Northeast

Dependent Variable: LHP
Method: Least Squares
Sample: 1963 2010
Included observations: 48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.702400	0.115536	40.70074	0.0000
LGDP	0.959623	0.024341	39.42480	0.0000
I	0.003033	0.004830	0.627883	0.5334
VR	0.385057	0.035121	10.96386	0.0000
DUMMY	-0.110388	0.026945	-4.096782	0.0002

R-squared	0.995395	Mean dependent var	11.64881
Adjusted R-squared	0.994967	S.D. dependent var	0.976754
S.E. of regression	0.069293	Akaike info criterion	-2.402603
Sum squared resid	0.206467	Schwarz criterion	-2.207687
Log likelihood	62.66248	Hannan-Quinn criter.	-2.328944
F-statistic	2323.920	Durbin-Watson stat	1.779140
Prob(F-statistic)	0.000000		

Midwest

Dependent Variable: LHP
Method: Least Squares
Sample: 1963 2010
Included observations: 48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.894015	0.076286	64.15379	0.0000
LGDP	0.934250	0.012946	72.16276	0.0000
I	0.018221	0.003474	5.245443	0.0000
VR	0.022739	0.022406	1.014867	0.3157

R-squared	0.995391	Mean dependent var	11.41257
Adjusted R-squared	0.995076	S.D. dependent var	0.822488
S.E. of regression	0.057713	Akaike info criterion	-2.787004
Sum squared resid	0.146556	Schwarz criterion	-2.631070
Log likelihood	70.88809	Hannan-Quinn criter.	-2.728076
F-statistic	3167.218	Durbin-Watson stat	0.632958
Prob(F-statistic)	0.000000		

South

Dependent Variable: LHP
Method: Least Squares
Sample: 1963 2010
Included observations: 48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.590903	0.059174	94.48186	0.0000
LGDP	0.801934	0.010097	79.42525	0.0000
I	0.005188	0.003428	1.513211	0.1374
VR	0.023211	0.024257	0.956900	0.3438

R-squared	0.996459	Mean dependent var	11.31786
Adjusted R-squared	0.996218	S.D. dependent var	0.869669
S.E. of regression	0.053484	Akaike info criterion	-2.939231
Sum squared resid	0.125861	Schwarz criterion	-2.783297
Log likelihood	74.54154	Hannan-Quinn criter.	-2.880303
F-statistic	4127.665	Durbin-Watson stat	0.561905
Prob(F-statistic)	0.000000		

West

Dependent Variable: LHP
Method: Least Squares
Sample: 1963 2010
Included observations: 48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.801336	0.117634	49.31685	0.0000
LGDP	0.862707	0.015157	56.91644	0.0000
I	0.001962	0.006156	0.318746	0.7514
VR	-0.006157	0.040286	-0.152818	0.8792

R-squared	0.988417	Mean dependent var	11.55796
Adjusted R-squared	0.987627	S.D. dependent var	0.941678
S.E. of regression	0.104746	Akaike info criterion	-1.594902
Sum squared resid	0.482756	Schwarz criterion	-1.438969
Log likelihood	42.27765	Hannan-Quinn criter.	-1.535975
F-statistic	1251.548	Durbin-Watson stat	0.315946
Prob(F-statistic)	0.000000		

Appendix 2 – Data sets

Northeast

Midwest

Year	Housing prices	Vacancy rates	Interest rate	GDP		Housing prices	Vacancy rates	Interest rate	GDP
1963	22,300	0.9	4	163.603		$19,700	1.5	4	177.538
1964	21,800	1.0	4.19	175.254		$20,700	1.3	4.19	189.31
1965	22,900	0.9	4.28	188.129		$22,800	1.2	4.28	207.82
1966	25,200	0.9	4.93	204.392		$24,600	1.0	4.93	225.712
1967	27,700	1.4	5.07	217.099		$26,400	1.0	5.07	235.21
1968	30,100	0.8	5.64	234.502		$28,500	1.0	5.64	256.12
1969	33,400	0.8	6.67	252.895		$29,900	0.9	6.67	276.546
1970	32,800	0.8	7.35	266.21		$28,000	1.0	7.35	282.673
1971	34,400	0.7	6.16	283.123		$29,900	0.9	6.16	307.741
1972	35,700	0.5	6.21	306.264		$31,400	1.0	6.21	337.426
1973	40,600	0.6	6.85	332.546		$36,700	1.0	6.85	381.139
1974	43,700	0.8	7.56	356.671		$39,300	1.0	7.56	401.083
1975	47,000	1.0	7.99	380.961		$43,400	1.0	7.99	433.715
1976	50,000	1.0	7.61	408.484		$48,600	1.0	7.61	487.995
1977	54,800	0.9	7.42	449.607		$55,200	1.0	7.42	539.748
1978	63,000	0.8	8.41	500.904		$64,200	1.0	8.41	603.806
1979	71,500	0.9	9.43	550.053		$73,000	1.1	9.43	658.416
1980	80,300	1.1	11.43	595.723		$74,400	1.6	11.43	680.613
1981	88,500	1.2	13.92	660.543		$82,500	1.4	13.92	751.22
1982	88,600	1.0	13.01	704.733		$87,700	1.6	13.01	766.418
1983	96,200	1.0	11.1	768.334		$97,600	1.5	11.1	818.489
1984	107,400	0.9	12.46	860.792		$107,800	1.6	12.46	920.729
1985	121,900	0.9	10.62	928.297		$95,400	1.6	10.62	972.675
1986	151,300	1.6	7.67	999.328		$102,600	1.5	7.67	1022.745
1987	170,900	1.7	8.39	1084.937		$115,500	1.4	8.39	1077.353
1988	179,300	1.6	8.85	1188.485		$123,700	1.2	8.85	1160.863
1989	188,600	1.8	8.49	1244.146		$130,600	1.3	8.49	1232.763
1990	190,500	1.7	8.55	1292.866		$133,000	1.3	8.55	1282.962
1991	188,800	1.7	7.86	1318.23		$134,500	1.3	7.86	1327.805
1992	194,900	1.5	7.01	1379.385		$136,400	1.2	7.01	1421.203
1993	183,600	1.4	5.87	1429.961		$143,100	1.1	5.87	1485.89
1994	200,500	1.5	7.09	1496.098		$152,700	1.1	7.09	1621.489
1995	216,600	1.5	6.57	1578.204		$157,200	1.3	6.57	1695.929
1996	226,800	1.6	6.44	1668.029		$158,100	1.4	6.44	1794.227
1997	234,100	1.6	6.35	1779.12		$173,000	1.2	6.35	1904.125
1998	240,100	1.7	5.26	1749.325		$179,200	1.4	5.26	1906.728
1999	247,900	1.7	5.65	1830.468		$186,800	1.3	5.65	2003.574
2000	274,800	1.6	6.03	1930.479		$199,300	1.3	6.03	2102.344
2001	294,300	1.8	5.02	2050.884		$201,300	1.7	5.02	2216.015
2002	301,300	1.7	4.61	2133.816		$209,800	1.8	4.61	2255.738
2003	315,700	1.8	4.01	2187.176		$218,200	1.7	4.01	2338.873
2004	366,100	1.7	4.27	2258.627		$240,800	2.0	4.27	2439.792
2005	397,000	1.9	4.29	2382.451		$249,800	2.3	4.29	2557.649
2006	428,300	2.4	4.8	2516.619		$257,100	2.6	4.8	2657.256
2007	437,700	2.7	4.63	2670.543		$256,800	2.8	4.63	2752.574
2008	475,500	2.8	3.66	2800.942		$250,000	2.9	3.66	2874.877
2009	411,300	2.6	3.26	2860.684		$227,700	2.6	3.26	2904.741
2010	418,400	2.6	3.22	2828.875		$230,600	2.6	3.22	2853.869

South

West

Year	Housing prices	Vacancy rates	Interest rate	GDP		Housing prices	Vacancy rates	Interest rate	GDP
1963	$16,800	1.8	4	154.311		$20,800	2.0	4	111.559
1964	$18,100	1.8	4.19	166.493		$23,200	1.8	4.19	119.457
1965	$18,900	1.9	4.28	181.186		$23,200	1.9	4.28	127.563
1966	$20,200	1.7	4.93	198.278		$25,500	2.2	4.93	138.693
1967	$21,100	1.6	5.07	212.719		$26,100	2.1	5.07	146.756
1968	$23,600	1.4	5.64	235.148		$27,100	1.5	5.64	161.795
1969	$25,300	1.2	6.67	257.624		$27,400	1.2	6.67	175.837
1970	$24,000	1.2	7.35	277.331		$26,900	1.1	7.35	186.085
1971	$25,900	1.2	6.16	305.058		$28,000	1.2	6.16	201.248
1972	$28,500	1.1	6.21	342.525		$30,500	1.2	6.21	224.08
1973	$33,200	1.2	6.85	390.362		$35,300	1.4	6.85	252.454
1974	$36,800	1.5	7.56	434.436		$39,300	1.5	7.56	279.831
1975	$39,600	1.5	7.99	476.362		$44,300	1.5	7.99	311.191
1976	$43,800	1.6	7.61	536.796		$51,900	1.1	7.61	347.221
1977	$48,100	1.7	7.42	599.566		$60,700	0.9	7.42	397.482
1978	$55,600	1.3	8.41	680.294		$70,100	1.0	8.41	458.225
1979	$63,800	1.1	9.43	763.667		$82,000	1.3	9.43	517.283
1980	$69,100	1.3	11.43	857.251		$89,400	1.6	11.43	580.343
1981	$75,600	1.4	13.92	988.917		$95,800	1.7	13.92	656.638
1982	$78,300	1.6	13.01	1044.445		$92,600	1.9	13.01	695.773
1983	$83,000	1.7	11.1	1111.698		$97,200	1.8	11.1	746.852
1984	$86,000	2.1	12.46	1237.279		$109,400	2.0	12.46	834.87
1985	$88,900	2.2	10.62	1327.034		$111,800	2.1	10.62	898.659
1986	$95,300	2.1	7.67	1370.904		$116,100	1.8	7.67	947.837
1987	$106,600	2.0	8.39	1460.058		$134,600	1.8	8.39	1026.672
1988	$114,800	1.9	8.85	1589.372		$155,700	1.6	8.85	1114.283
1989	$123,100	2.1	8.49	1691.144		$173,900	1.4	8.49	1198.025
1990	$123,500	2.1	8.55	1790.697		$180,600	1.8	8.55	1285.223
1991	$123,000	2.2	7.86	1869.198		$176,400	1.7	7.86	1327.432
1992	$126,900	1.8	7.01	1979.93		$157,800	1.9	7.01	1387.125
1993	$133,600	1.6	5.87	2100.023		$161,900	1.4	5.87	1451.822
1994	$136,800	1.7	7.09	2255.783		$168,900	1.6	7.09	1539.475
1995	$142,000	1.7	6.57	2393.184		$169,800	1.7	6.57	1632.056
1996	$143,100	1.8	6.44	2544.451		$185,900	1.7	6.44	1743.258
1997	$151,400	1.9	6.35	2726.276		$198,200	1.8	6.35	1874.911
1998	$159,700	2.0	5.26	2717.743		$200,500	1.7	5.26	1868.505
1999	$173,000	2.0	5.65	2871.277		$221,700	1.7	5.65	1992.299
2000	$179,000	2.0	6.03	3064.407		$238,900	1.5	6.03	2158.64
2001	$185,700	2.1	5.02	3254.07		$250,000	1.6	5.02	2329.68
2002	$197,500	1.9	4.61	3409.122		$276,500	1.6	4.61	2383.696
2003	$208,900	2.2	4.01	3537.146		$306,800	1.6	4.01	2471.125
2004	$232,800	2.0	4.27	3726.13		$340,000	1.4	4.27	2602.638
2005	$249,200	2.1	4.29	4007.023		$388,700	1.4	4.29	2798.854
2006	$257,700	2.7	4.8	4308.377		$405,900	2.1	4.8	3028.114
2007	$269,800	3.1	4.63	4590.657		$403,700	2.5	4.63	3250.719
2008	$253,400	3.2	3.66	4820.008		$361,500	2.9	3.66	3426.319
2009	$241,200	2.9	3.26	4939.975		$321,600	2.6	3.26	3516.858
2010	$242,800	2.8	3.22	4855.259		$315,700	2.7	3.22	3442.059