

The Predictability of Residential Real Estate Returns

Khaled Alsabah
University of California San Diego

under the direction of
Prof. Walter Torous
Center for Real Estate
Massachusetts Institute of Technology

Summer Research Internship
June 2, 2014 – August 22, 2014

Kuwait-MIT Center for Natural Resources and the Environment

August 22, 2014

1)Table of Contents:

- 2) Abstract..... 3
- 3) Introduction..... 4
 - a. Real Estate Market Efficiency..... 5
 - b. Brief Overview of Real Estate Data..... 6
 - c. Forecasting Returns..... 7
- 4) Data..... 7
 - a. Gross Rents..... 7
 - b. Home Values..... 8
- 5) Methodology..... 9
- 6) Results..... 10
- 7) Discussion and Further Research..... 11
- 8) Acknowledgements..... 13
- 9) References..... 13

2) Abstract:

This paper tests whether the rent-to-price ratio has any predictive power in forecasting residential housing returns in the US housing market, more specifically in six Metropolitan Statistical Areas (Boston, Los Angeles, Miami, Milwaukee, New York and San Diego). This study constructs time series data for both the nominal gross rent and nominal house values by using the Decennial Census of Housing data as benchmarks for each decade and then interpolate the in between values by using Freddie Mac House Price Index and Rent of Primary Residence Index for houses values and rents respectively. This study uses Ordinary Least Square method with Newey-West standard error to regress the results for different time horizons from 6 month up to 4 year intervals. This author finds that predictability strength is stronger at longer horizons. Also the 4 year horizon results indicate that for all the cities, except Boston, the t-statistic is so low that we cannot assume the valuation ratio of rent over price has any predictability power. On the other hand the semiannual horizon results indicate that we can assume predictability for all areas except Miami.

3) Introduction:

Residential real estate is an important asset for any country's economy. Equity in owner occupied housing is a large portion of national wealth, and any real appreciation in the value of single-family homes is likely to make a significant contribution to national saving. To gain a bit of insight about the significance of this sector let us have a closer look at the US market. The housing market's total value in the US at the end of 2013 was \$25 trillion which is equivalent to the stock market value in the same period. All home buyers view their decision to some extent as an investment decision (Case and Shiller, 1987). All of these facts strengthen the necessity of conducting research on the predictability of housing returns.

Although there is substantial amount of evidence of in-sample predictability of stock returns based on dividend yield ratio (see, among others, Fama and French (1988), Campbell and Shiller (1998), Choudhury (2003)), little is known about the predictability of residential real estate returns. A fact that has been taken into account when formulating the methodology for this paper is that predictability of dividend yields in the majority of papers (for example, Campbell 1991, Cochrane 1992) reveal that predictability is stronger at longer horizons relative to shorter horizon. Based on this fact this author have decided to test the predictability power by using different time intervals. Furthermore, a relatively small amount of research is available that test whether rent/price ratio provide forecasting power for commercial properties returns. From the few research papers that investigated this issue; the Plazzi, Torous, and Valkanov (2010) paper concludes that higher cap rates, which is another way of saying rent/price ratio, predicts higher (in-sample) return for apartment buildings as well as industrial and retail properties. They conclude that cap rates does not predict returns for office buildings. This paper will assess the in-sample predictability power of housing returns based on two different intervals, 6 month interval and a 4 year interval, by using the rent-to-price ratio as an independent variable.

The testable hypothesis of this paper is to confirm rent/price ratio has a similar predictability power for house returns as dividend yield for stock returns.

Before moving forward let us point out that the main intuition for using ratios to predict future returns lays in the realistic assumption that the variables used to form the ratio are co-integrated in logs (Engle and Granger (1987)). To put it into perspective lets consider the log rent-price ratio. If the logs of rents and prices are co-integrated, then the log ratio will always revert to the mean whenever it strays away.

For example if at some point the ratio is lower than the unconditional mean, then this implies that either expected growth rent will be high or expected returns will be low, or a combination of both (Plazzi, Torous, Valkanov (2010)).

A) Real Estate Market Efficiency:

Based on the huge amount of literature about dividend yield predictability and dividend theory, this author is enticed to believe that the rent/price ratio will have similar predictability power. As mentioned earlier, the rent/price ratio showed predictability power for apartment buildings, industrial and retail properties (Plazzi, Torous, Valkanov (2010)). At the same time this author has some concerns on whether the predictability power of similar valuation ratios, dividend yield and rent/price ratio, can be carried over to different markets, to be more specific from the more efficient stock market to the more constrained real estate market. The cap rate did not have any predictability power for offices (Plazzi, Torous, Valkanov (2010)). The real estate market is different from the stock market in many aspects that makes the predictability test more interesting and challenging. The real estate market is characterized by extreme heterogeneity that creates limitations that other financial markets do not face. For example, relatively speaking, real estate market participants encounter large transaction cost, illiquidity and are restricted from short selling to name a few of these market frictions. These large frictions suggest that the real estate market might be less efficient than other financial markets such as the stock market. This might imply that predicting housing returns should be an easier task given this inefficiency. It should be noted that being able to predict future returns does not necessarily imply that the market is inefficient. According to rational asset pricing theory, stock return predictability can result from exposure to time-varying aggregate risk by successfully capturing this time-varying aggregate risk premium, a forecasting model can remain successful over time.

According to a controversial study by Mankiw and Weil (1989), long trends in house prices can be predicted due to inefficiency. They argue that the rise of demand caused by the postwar baby boom could have been easily predicted 20 years in advance. Their empirical results show that home prices increase simultaneously rather than in advance whenever such information is available. This implies that the real estate market is not semi-strong efficient because public information, such as the demographic of the population, wasn't incorporated in house prices. Moreover, Hamilton and Schwab (1985) conclude that households' fails to accurately incorporate past appreciation in their expectation hence their result

does not support the rational expectation hypothesis which states that on average, people can quite correctly predict future conditions and take actions accordingly.

B) Brief Overview of Real Estate Data:

A fundamental factor for the few research papers in this area is the scarcity of data available on residential real estate in respect to both duration and frequency. Furthermore, the data can contain considerable amount of error if not modified due to high transaction and information costs, asymmetric information, and differences in the quality of rental versus owner occupied housing. This means that the predictive results have to take into consideration the transaction costs which can be 6%, or even higher, of the property value (Ghysels, Plazzi, Torous, and Valkanov, 2012). It should be noted that this current study uses disaggregated data so that the results will not overlook the individual characteristics of some cities due to diversification.

As mentioned earlier, relative to stocks data, real estate data are scarce in both duration and frequency. The house prices data used in this paper is based on repeated price indices — Macromarkets LLC national house price index — for a number of reasons. Unlike median prices indices, repeated price indices do not ignore the quality differences in housing. Repeated prices indices are constructed by using information of homes that have been sold at least twice in a specific period. In addition unlike hedonic price indices, repeated price indices do not require many variables to construct, at the very least we need only the price changes and dates of individual property transaction. On the other hand, it has some drawbacks such as limiting our available data substantially because homes that are sold only once are disregarded. Also one might argue that houses that are repeatedly sold are not representative of the whole market. Another concern with the data being used is that our rent-to-price ratio uses the annual rents paid for rental units to estimate annual rents for owner-occupied units. There will be measurement error in the data on rent/price ratio because owned and rental units have different qualities. Furthermore, the average price of owner occupied house used in the denominator is the average of self-reported value presented by owners. Allowing self-reporting might cause some error in the data especially when owners tend to overestimate their property's value.

On the other hand, it should be noted that real estate markets offer a number of advantages relative to stock market when it comes to valuation ratio such as rent/price ratio, analogous to dividend yield. An advantage of the rent/price ratio over the dividend yield is that rents, as opposed to dividends, cannot

be manipulated because tenants are the ones who pay the rents as opposed to managers who issue the dividends themselves. Furthermore, there is extensive evidence that dividends are subjected to active smoothing (Leary and Michaely (2009)). Finally housing rents are directly affected by macroeconomic factors such as employment (Davenport (2003)).

C) A Brief Review of Asset Pricing:

Forecasting stock returns is a fascinating endeavor with a long history. This paper will examine the predictability of housing using tools that have proven to be successful in analyzing stock market returns. From an asset pricing perspective, the price of housing equals the expected net present value of future rents the home owner would gain by renting the unfurnished property, which includes the expected growth in rents. This fundamental present value relation implies that observed fluctuations in home prices should reflect variation in future rents, in future discount rates, or both. For example when rent to price ratios fell during the housing boom of 1997-2006, this implied that either the discount rate on housing had fallen, the expected growth rate of housing rents had increased, or combination of both.

4) Data Used In This Study:

In this paper, a semiannual time series was developed, starting from 1975 and ending in 2013, of the ratio of imputed gross rent of homes over the value of owner occupied housing using nominal values.

A) Gross rent:

The gross rent data includes the cost of utilities in the addition to the contract rent. This paper employ a two-step procedure, similar to the Davis, Lehnert and Martin (2007) method which was conducted for an aggregated level, to construct the price-rent ratio for each Metropolitan Statistical Area (MSA). The first step is to develop benchmarks for mean gross rents. The data for gross rents have been gathered from two main sources, DCH and AHS, to develop benchmark estimates for mean gross rent for each MSA. First this author have collected micro data for each MSA being studied from the Decennial Census of Housing (DCH) surveys, 1970-2000. It should be noted that the DCH surveys ceased to exist after its last issue on 2000 and benchmark data for post 2000 were constructed from American Housing Survey (AHS). Unlike DCH, AHS are conducted non-uniformly between each survey. For example, San Diego surveys have been conducted in 1975, 1978, 1982, 1987, 1991, 1994, 2002 and 2011. Although the time

interval between each survey is non-uniform, in every decade there is at least one survey which satisfies the minimum requirement to construct the proposed estimated benchmark. Furthermore, the surveys conducted by DCH and AHS only issue the median gross rent of each metropolitan area except the 1990 DCH survey which issued the mean gross rent as well. Given the estimated benchmarks are based on averages, this author calculated the mean gross rent using the detailed housing characteristic by allocating the appropriate weight to each rental interval based on the number of units found in each segment.

After forming the required benchmarks, the second step is to interpolate gross rents between each benchmark and extrapolate beyond the last benchmark using Bureau of Labor Statistics (BLS) rent of primary residence (RPR) index. This author have adjusted the observed semiannual changes in the RPR index so that it passes through the estimated benchmarks levels.

Although BLS offers owners' equivalent rent of residence (OERR) index, this author preferred to use rent of primary residence for a number of reasons. First, RPR provide a longer time series relative to OERR which is vital for this study. To put it into context, RPR offer indices starting from 1967 while OERR starts at 1982, hence increasing the database pool by 38%. Second, RPR provide an accurate measurement of rent paid by tenants while OERR is based on owner expectations on how much his/her home would generate per month. Finally, the estimated benchmarks used for gross rents are derived from renter occupied units which is why using RPR seems better fit with the existed data.

The Bureau of Labor Statistics (BLS) data only publishes the rent indices for 23 metropolitan areas which is the reason for limiting our data to these MSA's. In addition, BLS data are repeated at semi-annual frequency for most of the years and is why this author uses this frequency to conduct his analysis. Furthermore the data of this paper starts from 1975 because this is when Freddie Mac House Price Index (FMHPI) started to publish its data.

B) Home Values:

The home values includes both the land and structure cost. The home values data gathered are from Lincoln Institute of Land Policy (LILP). LILP database included 46 metropolitan areas spanning from 1984 to 2013. Their data was estimated by using AHS to develop the benchmarks and then extrapolated forward and backward by using metro-area CMHPI and Case-Shiller-Weiss house price indexes. It should be noted that the CMHPI have ceased to be published and was replaced by FMHPI at the start of 2011.

This paper uses Freddie Mac House Price Index (FMHPI) to extrapolate backward of the already existed LILP data and estimate the missing home values of pre 1984. Also, this paper have used the 1984 home value as its benchmark to extrapolate backward. In addition, this author converted FHMPI monthly indices into semiannual so that it agrees with our frequency interval.

Finally, there are some concerns related to the data used in this study. First, by using repeated sales indices such as FMHPI, our home values might be biased upward. According to several researchers (Case, Pollakaowski and Watcher (1997), Gatzlaff and Haurin (1997)) homes that are traded frequently tend to have greater price appreciation. Second, this paper uses the rental price of tenant occupied housing as a substitute to the rental price of owner-occupied housing. If the rent trends in owner occupied housing is not captured by the rents of tenant-occupied housing then our analysis can be affected. Lastly, researchers (Crone, Nakamura and Voith (2004), Lebow and Rudd (2003)) have stated that BLS does not adjust for aging effects properly and should increase its growth rate. Taken these considerations, this author expect that the ratio calculated in this paper is lower than actual because of the upward bias of home values and downward bias of rent. It should be noted that lower ratios should not affect our analysis, if the trends of the ratio over time stayed the same which is what we expect.

5) Methodology:

This paper will be using Ordinary Least Square (OLS) time series regression to test for the predictability power of the rent/price ratio in forecasting housing returns. For longer horizon, s , we will have to adjust the standard errors due to the fact of overlapping observations which will generate autocorrelation. Due to autocorrelation and heteroscedasticity, OLS-t values usually will be larger than the actual value. As a result, that upward bias might lead to a type-I error where the true hypothesis of no predictability is incorrectly rejected. To rectify the autocorrelation complication, we will adjust the errors with Newey-West (1987) standard errors, or the so-called heteroscedasticity and autocorrelation - consistent (HAC) standard error. This paper will be testing two time horizons (1) single semiannual frequency and (2) four years horizon as shown in the following equations:

$$(1) R_{t+1} - R_t = b_0 + b_1(r_t / p_t) + e_t$$

$$(2) R_{t+8} - R_t = b_0 + b_1(r_t / p_t) + e_t$$

b_0 : constant coefficient (alpha)

b_1 : rent/price ratio coefficient (beta)

e_t : error term

r_t : gross rent of residential housing in nominal dollars

p_t : value of residential housing in nominal dollars

R_t : return of housing, calculated as $R_t = (r_t + p_t) / p_{t-1} - 1$

6) Results:

Table 1: Semiannual Frequency (1975-2013 period)

City	b_0	b_1	R ² %	t-Statistics
Boston	-0.0183435	2.430437	19.99	5.19
Los Angeles	-0.0203845	2.680099	8.84	2.32
Miami	0.0463047	0.3923324	0.25	0.3
Milwaukee	-0.0245357	1.950052	15.79	4.7
New York	-0.026063	2.017769	9.6	3.17
San Diego	-0.0127355	2.335629	6.37	2.04

Table 1 present the results of equation (1) time series regression. The first two columns represent the coefficients, alpha and beta respectively. The third column represent R-squared in percentile which indicates the fraction of the variation in my dependent variable, R_t , that is accounted for through the independent variable, r_t/p_t . The fourth column represent the OLS t-statistic. Data frequency is semiannual.

Table 2: Time horizon 4 years (1975-2013 period)

City	b_0	b_1	R ² %	t-Statistics	N-W t-Statistics
Boston	-0.0652708	3.615501	43.82	7.13	2.3
Los Angeles	-0.0826028	4.288653	26.12	4.33	0.34
Miami	-0.1035428	4.229623	28.63	4.28	-1.16
Milwaukee	-0.0660536	2.739103	41.84	7.1	1.4
New York	-0.0854596	4.630807	50.66	8.45	0.89
San Diego	-0.0841535	4.29759	25.2	3.61	0.06

Table 2 present the results of equation (2) time series regression. The first two columns represent the coefficients, alpha and beta respectively. The third column represent R-squared in percentile which indicates the fraction of the variation in my dependent variable, R_t , that is accounted for through the independent variable, r_t/p_t . The fourth column represent the OLS t-statistic while the fifth column represent the Newey-West t-statistic.

7) Discussion and Future Research:

Similar to numerous research papers predicting stock returns, the results confirm that housing returns predictability in terms of magnitude and R-squared do increase with longer horizon. This suggests that longer horizon has stronger predictability, but examining the N-W t-statistics provide us with a different interpretation. For the 4 year horizon, all of the N-W t-statistics are low that we cannot reject the null hypothesis of no predictability except for Boston case. On the other hand, the semiannual horizon has more significant coefficient due to the larger t-statistics which allows us to reject the null hypothesis of no predictability for all areas except Miami.

Future work on residential housing predictability would include more MSA, than the current paper. To be specific future papers will add 17 more MSA because BLS publish the data for 23 MSA. With more disaggregated data, we will be able to reach a more robust conclusion about housing return predictability. Furthermore, predictability in housing might be affected by the boom the housing market experienced during 1997-2007. Future papers will examine the data over three period of times 1975-1996, 1997-2007 and 2008-2013 equivalent to pre-boom, boom and post boom. In addition we will test more time horizons such as a 2 year and a 3 year horizon, in case these intermediate horizons might be able to show us something that our 4 year horizon have missed.

The rent to price ratio might be able to predict rent growth in addition to housing prices, given the co-integrated relationship between rent and price. Future work would include the predictability of rent growth in addition to housing returns. Finally, a thorough examination of macroeconomic factors might help clarify the discrepancy why some MSAs have predictability powers while others do not. Some variable in this regard that might play an important role would be population density and land use regulation. According to the Plazzi, Torous and Valkanov (2010) paper, commercial real estate is affected by these two factors. They found evidence that return predictability is stronger in locations with lower population density and less stringent land use restrictions. On the other hand, rent growth predictability is more likely observed in locations characterized by higher population density and more severe land use restrictions. There is no reason to believe that these trends cannot be carried over from commercial real estate toward residential real estate.

8) Acknowledgment:

First I highly appreciate Prof. Walter Torous guidance and supervision which was vital in proceeding with this paper. Also a special thanks to both Dr. Murad Abu-Khalaf and Laura Guild for facilitating my transition into MIT. I would like to thank Kuwait Foundation for the Advancement of Sciences (KFAS) for sponsoring me. Finally I would like to thank Kuwait-MIT Center for Natural Resources (CNRE) for providing this great learning opportunity.

9) References:

- [1] Himmelberg, C., C. Mayer and T. Sinai (2005), "Assessing High House Prices: Bubbles, Fundamentals, and Misperception. "
- [2] Davis, Morris A. and Jonathan Heathcote, (2007), "The Price and Quantity of Residential Land in the United States," *Journal of Monetary Economics*, vol. 54 (8), p. 2595-2620
- [3] Gallin, J (2008), "The Long-Run Relationship Between House prices and Rents", *Real Estate Economics* V36 4:pp. 635-658
- [4] Fama, E.F. and K.R. French (1988), "Dividend Yields And Expected Stock Returns," *Journal of Financial Economics* 22 (1988) 3-25
- [5] Hodrick, R.J. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *The Review of Financial Studies*
- [6] Plazzi, A, W. Torous and R. Valkanov (2010), "Expected Returns and Expected Growth in Rents of Commercial Real Estate."
- [7] M.A Davis, A. Lehnert and R.F. Martin (2007), "The Rent-Price Ratio for the Aggregate Stock of Owner-Occupied Housing."
- [8] N.G. Mankiw and D.N. Weil (1989), "The Baby Boom, The Baby Bust, And The Housing Market," *Regional Science and Urban Economics* 19 (1989) 235-258.
- [9] J. Y. Campbell and R.J. Shiller (1988), "Stock Prices, Earnings, and Expected Dividends," *The Journal of Finance*, Vol. 43, No. 3, pp. 661-676.
- [10] M. Lettau and S.V. Nieuwerburgh (2008), "Reconciling the Return Predictability Evidence," *The Review of Financial Studies* v21, n4 pp. 1607-1652.
- [11] E. Ghysels, A. Plazzi, W. Torous and R. Valkanov (2012), "Forecasting Real Estate Prices"

- [12] K.E. Case and R.J. Shiller (1987), "Prices of Single Family Homes Since 1970: New Indexes For Four Cities," Cowles Foundation Discussion Paper no. 851.
- [13] N.K Choudhury (2003), "Does the Dividend Yield Predict International Equity Returns."
- [14] J.Y. Cambell (1991), "A Variance Decomposition For Stock Returns" NBER Working Paper No. 3246
- [15] J.H Cochrane (1992), "Explaining the Variance of Price-Dividend Ratios," Review of Financial Studies, Society for Financial Studies, Vol5(2), pp.243-80.
- [16] R.F. Engle and C.W.J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing," Econometrica, Vol. 55, No. 2, pp.251-276.
- [17] B. W. Hamilton and R. M. Schwab (1983), "Expected Appreciation in Urban Housing Markets," Journal of Urban Economics 18, pp. 103-118.
- [18] W. K. Newey and K. D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity And Autocorrelation Consistent Covariance Matrix," Econometrica, Vol 55, No. 3, pp. 703-708.
- [19] T. M. Crone, L. I. Nakamura and R. P. Voith (2004), "Hedonic estimates of the cost of housing services: rental and owner-occupied units," Working Papers 04-22, Federal Reserve Bank of Philadelphia.
- [20] D. E. Lebow and J.B. Rudd (2003), "Measurement Error in the Consumer Price Index: Where Do We Stand?," Journal of Economic Literature Vol. XLI pp. 159-201
- [21] D. H. Gatzlaff and D. R. Haurin (1997), "Sample Selection and Biases in Local House Value Indices," Journal of Urban Economics, Volume 43, Issue 2, March 1998, Pages 199-222.
- [22] J. L. Davenport (2003), "The Effect of Supply and Demand Factors on the Affordability of Rental Housing."
- [23] A. Goyal and I. Welch (2001), "Predicting the Equity Premium with Dividend Ratios," Management Science/Vol. 49, No. 5 May 2003
- [24] M. T. Leary and R. Michaely (2009), "Determinants of Dividend Smoothing: Empirical Evidence."