Performance Persistence in Equity Real Estate Returns

by

Michael S. Young
Vice President and Director of Quantitative Research
The RREEF Funds
101 California Street, San Francisco, CA 94111
phone: 415-781-3300 / fax: 415-781-2229 / e-mail: MYoung@RREEF.com

and

Richard A. Graff
Principal
Electrum Partners
400 North Michigan Avenue, Suite 415, Chicago, IL 60611
phone: 312-923-8000 / fax: 312-923-8023

published in

Real Estate Finance
Vol. 14, No. 1, Spring 1997, pp. 37-42

Performance Persistence in Equity Real Estate Returns

by
Michael S. Young and Richard A. Graff

Most real estate researchers have assumed that equity real estate returns are independent and normally distributed across time. Most real estate researchers appear to have been mistaken.

In the real estate asset class, recent empirical studies by the authors and others have questioned the applicability of Modern Portfolio Theory (MPT) or the Efficient Markets Hypothesis (EMH) in their current forms.[1] Our findings here are another in that vein.

We show that serial dependence of total annual returns in the National Council of Real Estate Investment Fiduciaries (NCREIF) data base is statistically significant in the first and fourth quartiles between 1978 and 1994. More precisely, performance is persistent: superior performance is generally followed by continued superior performance, and inferior performance is generally followed by continued inferior performance. In contrast, there is virtually no evidence to support serial dependence in the second or third quartiles.

The empirical rejection of serial independence among real estate returns calls into question the conclusions of research based upon models that incorporate the assumption of serial independence, i.e., most of the extant real estate performance-related research. Prices and returns in an efficient market should appear to fluctuate randomly; there should not be trends or runs in the data. However, persistent runs are exactly what we found in a particular portion of real estate returns data. If MPT or EMH are valid models for equity real estate, our finding of performance persistence at the extremes of the risk distribution should not have been statistically significant.

Jensen [1969] investigated the ability of stock mutual funds (and their managers) to outperform the market in succeeding periods by dividing the performance of funds into two groups, superior and inferior. He found no evidence of serial dependence in mutual fund performance. Young & Graff [1995] found that real estate asset-specific risk was, more often than not, negatively skewed, which leads us to conduct a nonparametric test similar to Jensen's, but with more groupings. We use performance quartiles to contrast the extremes, the first and fourth quartiles, with the center, the second and third quartiles.

Testing Runs for Persistence

We examine the serial persistence of annual total returns in the appraisal-based NCREIF Property Index between 1978 and 1994 grouped by one of five property types within Metropolitan Statistical Areas (MSAs).[2] To preserve anonymity, NCREIF combines data within

---

[1] Substantial departures from Gaussian normal return distributions have been noted by Liu, et al. [1992], Myer & Webb [1990 & 1993], and Young & Graff [1995].

[2] As of December 31, 1995, the NCREIF Property Index consisted of 2,322 properties having an estimated aggregate market value of approximately $47.8 billion. By various estimates this represents about one-third of all equity real estate owned by domestic public and private pension plans. For this study, properties having fewer than four quarters of data for a year were eliminated.

each MSA so that no fewer than four properties occupy any MSA/property-type cell.[3] The total number of MSAs ranged from 8 in 1978 to 44 in 1991; the total number of MSA/property type cells ranged from 9 in 1978 to 134 in 1992. Except for Apartment returns that begin in 1985, all other property-type returns span the entire 1978 to 1994 period.

For each year, we arrange annual total return performance into quartiles—cross-sectional groupings, and record the quartile rank for the year following a uniform within-quartile run of from 1 to 6 years—sequential groupings. Successful performance persistence is defined as an identical quartile rank in the subsequent year. A description of the procedure for testing serial persistence is shown in the Appendix.

Because the cross-sectional data are divided into quartiles, the theoretical probability of success is 25% if quartile rankings are serially independent. Because each MSA/property-type return quartile is independent across time, the probability of returns remaining in the same quartile from one period to the next is 25%.[4] We consider statistically significant departures from 25% evidence of serially dependent performance persistence.

For each year from 1978 to 1994, the total returns for each MSA/property-type cell were designated as being in one of four quartiles with the first quartile having the highest total returns and the fourth quartile having the lowest total returns.[5] The data were also grouped by property type to examine serial dependence along this dimension.

Confidence Interval Estimation for Discrete Samples

To ascertain whether quartile performance is serially dependent, we calculated confidence intervals for the binomial distribution under the assumption that the probability of repeating quartile performance is 25%. The sample statistic is the percent of sample MSA/property-type returns for which the quartile rank in the subsequent sample period equals the quartile rank during the immediately preceding series of sample periods. The central question is whether or not the sample statistic is statistically distinct from 25%.

Because the binomial distribution is discrete, the sample statistic can only assume a finite number of potential values between 0 and 1. Smooth probability distributions by contrast have a positive probability that a sample value for the statistic can equal one of the end points of a confidence interval. To avoid confusion about whether or not the sample value is within the confidence interval, the left and right end points of the confidence intervals reported in the

---

[3] NCREIF policy prohibits publication of individual property returns in order to preserve the privacy of contributing member firms and their clients. The current rule is that performance data be aggregated such that a minimum of four properties from at least two investment managers occupy the smallest statistical data cell. Other MSA/property type groupings may have existed but the "two-manager" rule prevented their inclusion in this study. Also, aggregating property data into a single return for each MSA/property type cell produces a value-weighted total return rather than a mean of the total returns of each property computed separately. This value-weighting is fortuitous in two respects. First, investor portfolio returns are typically value-weighted so this approach is consistent with the view of most investors. Second, the results would have been far less interesting if they could have been attributed to the contributions of small, relatively low-valued properties.

[4] This statement is less restrictive than the assertion that the returns of MSAs are independent across time.

[5] In a separate series of tests not reported here, we arranged the cross-sectional data by quartiles for each property-type and looked for performance persistence separately for each type. The results were essentially the same but somewhat noisier in the first and fourth quartiles.

exhibits lie midway between two possible sample values of the binomial distribution. Each sample value is unambiguously either inside or outside each confidence interval.

Since the number of cross-sectional returns is seldom evenly divisible by four, the number of data points in each quartile are not always equal. Because of the way we set the quartile breaks however, any bias in quartile groupings is toward the center of the distribution, toward more second and third quartile designations.

Persistence in the Extremes, But Not in the Center

We find that real estate equity performance persistence is statistically significant in the extremes of the cross-sectional distribution of NCREIF returns grouped by property type within MSAs for the years 1978 to 1994. By contrast, performance persistence is virtually undetectable in the central 50-percentile range of cross-sectional returns. The power of the statistical tests in the combined extreme first and fourth quartiles is concrete evidence of systematic behavior in investment risk, a finding that invalidates current formulations of portfolio theory and practice within the real estate asset class that rely upon assumptions of normal distributions and uniform risk across time.

Exhibit 1 shows that persistence of total annual returns within the combined first and fourth quartiles is statistically significant for each property type and for all property types taken together following runs of 1 year for Retail and R&D properties, following runs of 1 and 2 years for Warehouse properties, and following runs of 1, 2, and 3 years for Office and Apartment properties. Combining all property types we find that statistical significance of performance persistence occurs following runs of 1, 2, 3, 4, and 5 years. Moreover, among the 15 statistically significant findings shown in Exhibit 1, five are statistically significant at the 99.99999% level, and ten are statistically significant at the 99.9% level.

The results in the center of the return distribution—results in the combined second and third quartiles—stand in sharp contrast to the results at the extremes. Exhibit 2 shows that Warehouse properties had statistically significant return persistence following a 1-year run at the 99% confidence level and that the combined property data had statistically significant return persistence following 1-year and 2-year runs. All other runs by property type or among all property types combined failed to show performance persistence distinguishable from the theoretical expected value of 25%. Thus, randomness in the center of the distribution of returns is prevalent, and differs substantially from the serially dependent results at the high and low extremes.

Exhibit 3 shows the results of combining all performance data from each of the four quartiles. The strong evidence of performance persistence is not diluted by combining the weak middle of the distribution with the strong extremes. Indeed, the statistical significance of persistence following a 3-year run improves from the 99.9% level in the combined first and fourth quartile results to the 99.99999% level when all quartiles are taken together.

How Might Investors Use this Information about Persistence?

This empirical examination demonstrates conclusively that total returns of properties grouped at the MSA/property-type level within the NCREIF data base exhibit serial dependence at the extremes of the distribution. This result is at odds with the prevailing assumption about real estate risk, and calls into question current beliefs and portfolio construction applications. It

follows that the conclusions of research based upon models that incorporate the assumption of serial independence are suspect.

At first glance, performance persistence suggests a simple investment decision rule: hold your extreme winners; sell your extreme losers. For investments in the center of the MSA/property-type performance distribution, a more detailed examination of the risk/reward characteristics of each property is warranted before making portfolio decisions.

Rather than suggest investment decision rules, we believe that the most useful application of these results is in "filter" or "screening" rules: rules that indicate what and where to avoid, and what and where to seek new investment opportunities. While asset selection should dominate allocation in general, knowledge of persistence may be used to increase the probability of success in the extreme upper quartile of MSA/property-type group and to avoid inferior performance in the extreme lower quartile.

If the knowledge of performance persistence suggested here becomes widely known by investors, we might expect any value created by knowledge of persistence to dissipate quickly as prices are bid up or down to reflect active buying in known superior MSA/property-type groups or to reflect active retreat from known inferior MSA/property-type groups. For an investment strategy based on persistence to succeed, there must be widespread disbelief in its efficacy. Alternatively, investors could believe in persistence but face restrictive investment constraints that inhibit acting upon its knowledge.

Based on the information presented here, we cannot reject the possibility that some of the observed persistence in the extreme quartiles could be due to the widely-discussed phenomenon of appraisal smoothing. This possibility will be addressed in a future article.

Appendix

Counting Runs and Persistence

The manner by which we count runs and determine persistence can be shown by way of example from the actual sample data. The following table is an excerpt from the sample data for the office property type in five MSAs covering the years 1986 to 1994, the last full year of our sample.

Quartile Rank for Office Properties in a Sample of Five MSAs
1986 to 1994

MSA	1986	1987	1988	1989	1990	1991	1992	1993	1994
Anaheim	4	2	2	3	4	3	4	4	1
Atlanta	3	4	4	3	4	4	2	3	1
Austin						4	4	1	2
Boston	2	2	1	4	3	4	4	3	4
Charlotte				1	3	3	3	3	2

Reading across, we find that Anaheim has no runs of first quartile rank (we cannot count the 1994 rank of 1 as a run because the "end point" must be reserved to examine whether or not persistence exists for the year immediately preceding the "end point," i.e., for the year 1993). Anaheim has two runs of length 1 of second quartile rank: years 1987 and 1988; two runs of length 1 of third quartile rank: years 1989 and 1991; and four runs of length 1 of fourth quartile rank: years 1986, 1990, 1992, and 1993.

Staying with Anaheim, we find that there are only two runs of length 2: one run of second quartile performance in the years 1987 and 1988, and one run of fourth quartile performance in the years 1992 and 1993. Runs of length 2 give rise to persistence following a run of length 1, what we have termed a "success." In the Anaheim data, we find that a second quartile run of length 1 in 1987 was followed by persistence within the second quartile in 1988. Similarly, a fourth quartile run of length 1 in 1992 was followed by persistence within the fourth quartile in 1993.

The Charlotte office returns demonstrate imbedded runs. There are four third-quartile runs of length 1 for the years 1990, 1991, 1992, and 1993. The first three of these years show persistence in the third quartile because the following year also exhibits a third quartile rank. Also, there are three third-quartile runs of length 2 for the years 1990 and 1991, 1991 and 1992, and 1992 and 1993. Only the first two pairs of years show persistence in the following year. Next, there are two third-quartile runs of length 3 for the years 1990, 1991, and 1992; and 1991, 1992, and 1993. Only the first run of three years shows persistence in the following year, i.e., continuation in the third quartile in 1993. Finally, there is one third-quartile run of length 4, but no persistence in the year immediately following.

We follow this procedure of counting runs and successful persistence for each MSA/property-type data sample. The results for the small set shown above would be tallied as follows:

	First Quartile Performance			Second Quartile Performance		
MSA	Length of Run	No. of Samples	No. of Successes	Length of Run	No. of Samples	No. of Successes
Anaheim	1			1	2	1
				2	1	0
Atlanta	1			1	1	0
Austin	1	1	0	1		
Boston	1	1	0	1	2	1
				2	1	0
Charlotte	1	1	0	1		

	Third Quartile Performance			Fourth Quartile Performance		
MSA	Length of Run	No. of Samples	No. of Successes	Length of Run	No. of Samples	No. of Successes
Anaheim	1	2	0	1	4	1
				2	1	0
Atlanta	1	3	0	1	4	2
				2	2	0
Austin	1			1	2	1
				2	1	0
Boston	1	2	0	1	3	1
				2	1	0
Charlotte	1	4	3	1		
	2	3	2			
	3	2	1			
	4	1	0			

References

Jensen, Michael C. “Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios.” *Journal of Business*, April 1969, pp. 167-247.

Liu, Crocker H., David J. Hartzell, and Terry V. Grissom. “The Role of CO-Skewness in the Pricing of Real Estate.” *Journal of Real Estate Finance and Economics*, 1992, 5, pp. 299-319.

Myer, F.C. Neil and James R. Webb. “Are Commercial Real Estate Returns Normally Distributed?” working paper, Cleveland State University, 1990.

Myer, F.C. Neil and James R. Webb. “Return Properties of Equity REITs, Common Stocks, and Commercial Real Estate: A Comparison.” *Journal of Real Estate Research*, 1993, 8, pp. 87-106.

Young, Michael S. and Richard A. Graff. “Real Estate is Not Normal: A Fresh Look at Real Estate Return Distributions.” *Journal of Real Estate Finance and Economics*, 1995, 10, pp. 225-259.

Exhibit 1
Persistence in the Combined First and Fourth Quartiles of NCREIF Returns
1978 to 1994

Property Type	Length of Run	No. of Samples	No. of Successes	Percent of Successes	95% Confidence Interval
Office	1	173	93	53.8****	(18.2, 32.1)
	2	81	44	54.3****	(15.4, 35.2)
	3	37	24	64.9***	(9.5, 41.9)
	4	19	13	68.4	(2.6, 50.0)
	5	12	7	58.3	[0.0, 54.2)
	6	6	3	50.0	[0.0, 75.0)
Retail	1	90	35	38.9**	(16.1, 35.0)
	2	31	13	41.9	(8.1, 43.5)
	3	11	3	27.3	[0.0, 59.1)
	4	3	1	33.3	[0.0, 83.3)
	5	1	0	0.0	[0.0, 100.0]
Warehouse	1	139	56	40.3***	(17.6, 32.7)
	2	52	23	44.2**	(12.5, 37.5)
	3	22	10	45.5	(6.8, 47.7)
	4	10	5	50.0	[0.0, 55.0)
	5	5	2	40.0	[0.0, 70.0)
	6	2	0	0.0	[0.0, 100.0]
R&D	1	82	34	41.5**	(15.2, 34.8)
	2	31	9	29.0	(8.1, 43.5)
	3	9	2	22.2	[0.0, 61.1)
	4	2	0	0.0	[0.0, 100.0]
Apartment	1	55	33	60.0****	(13.6, 37.3)
	2	23	15	65.2***	(6.5, 45.7)
	3	9	6	66.7*	[0.0, 61.1)
	4	3	2	66.7	[0.0, 83.3)
All Samples Combined	1	539	251	46.6****	(21.2, 28.8)
	2	218	104	47.7****	(19.0, 31.0)
	3	88	45	51.1***	(15.3, 34.7)
	4	37	21	56.8***	(9.5, 41.9)
	5	18	9	50.0*	(2.8, 47.2)
	6	8	3	37.5	[0.0, 68.8)

* Statistically distinct from 25.0% with 95% confidence
** Statistically distinct from 25.0% with 99% confidence
*** Statistically distinct from 25.0% with 99.9% confidence
**** Statistically distinct from 25.0% with 99.99999% confidence

Exhibit 2
Lack of Persistence in the Combined Second and Third Quartiles of NCREIF Returns
1978 to 1994

Property Type	Length of Run	No. of Samples	No. of Successes	Percent of Successes	95% Confidence Interval
Office	1	145	38	26.2	(17.6, 32.8)
	2	32	9	28.1	(7.8, 42.2
	3	8	4	50.0	[0.0, 68.8)
	4	3	1	33.3	[0.0, 83.3)
Retail	1	106	28	26.4	(16.5, 33.5)
	2	22	6	27.3	(6.8, 47.7)
	3	5	2	40.0	[0.0, 70.0)
	4	1	0	0.0	[0.0, 100.0]
Warehouse	1	180	61	33.9**	(18.6, 31.9)
	2	53	19	35.8	(12.3, 38.7)
	3	16	4	25.0	(3.1, 53.1)
	4	3	0	0.0	[0.0, 83.3)
R&D	1	94	30	31.9	(16.5, 34.6)
	2	25	9	36.0	(6.0, 46.0)
	3	8	5	62.5	[0.0, 68.8)
	4	5	3	60.0	[0.0, 70.0)
	5	3	1	33.3	[0.0, 83.3)
	6	1	0	0.0	[0.0, 100.0]
Apartment	1	32	12	37.5	(7.8, 42.2)
	2	8	4	50.0	[0.0, 68.8)
	3	2	0	0.0	[0.0, 100.0]
All Samples Combined	1	557	169	30.3**	(21.3, 28.8)
	2	140	47	33.6*	(17.5, 32.5)
	3	39	15	38.5	(11.5, 39.7)
	4	12	4	33.3	[0.0, 54.2)
	5	3	1	33.3	[0.0, 83.3)
	6	1	0	0.0	[0.0, 100.0]

* Statistically distinct from 25.0% with 95% confidence
** Statistically distinct from 25.0% with 99% confidence
*** Statistically distinct from 25.0% with 99.9% confidence
**** Statistically distinct from 25.0% with 99.99999% confidence

Exhibit 3
Persistence of NCREIF Returns in All Quartiles Combined
1978 to 1994

Property Type	Length of Run	No. of Samples	No. of Successes	Percent of Successes	95% Confidence Interval
All Samples Combined	1	1,096	420	38.3****	(22.4, 27.6)
	2	358	151	42.2****	(20.5, 29.7)
	3	127	60	47.2****	(16.9, 33.5)
	4	49	25	51.0***	(13.3, 37.8)
	5	21	10	47.6*	(7.1, 45.2
	6	9	3	33.3	[0.0, 61.1)

* Statistically distinct from 25.0% with 95% confidence
** Statistically distinct from 25.0% with 99% confidence
*** Statistically distinct from 25.0% with 99.9% confidence
**** Statistically distinct from 25.0% with 99.99999% confidence