Cointegration in Return Series and Its Effect on Short-Term Prediction

by *Sorin A. Tuluca,* Fairleigh Dickinson University, Department of Finance, *Michael J. Seiler,* Hawaii Pacific University, Department of Finance, *F.C. Neil Myer,* Cleveland State University, Department of Finance and *James R. Webb*, Cleveland State University, Department of Finance.

This study investigates the implications of cointegration on a system of real and financial assets. The assets examined are: Treasury bills, long-term corporate bonds, large capitalization common stocks, and securitized and unsecuritized real estate. The five asset classes are shown to be cointegrated with two cointegrating vectors. Error correction augmented structural VAR models (VECM) are used to assess the improvement in prediction of asset returns in relation to unrestricted VAR models. The returns for unsecuritized real estate are forecast with improved accuracy while the other asset classes have similar properties regardless of which model is used. Finally, the implications of cointegration for investment decisions are discussed.

Introduction

There is considerable interest in the study of asset classes for investment diversification, portfolio allocation, and market timing strategies. However, the existing literature mainly investigates the risk and returns of asset classes, giving little or no consideration to the fact that returns are obtained from price indices by differencing. That several price indices may represent a cointegrated system of individual nonstationary series is not generally given much attention, even though it has been widely recognized in the last ten years that many financial and economic series are cointegrated. If price indices are cointegrated, it changes the way models of returns (the first difference of a price index) should be built and statistical conclusions drawn.

This study investigates five asset classes: short-term bonds, long-term bonds, common stocks, unsecuritized real estate, and securitized real estate as a cointegrated system. This has not been done before. Recognizing that returns are derived from wealth (price) indices, the cointegration among indices for the five asset classes is analyzed. Both long-term relationships among indices, and the possibility of short-term prediction of the returns on the five asset classes is analyzed.

The remainder of the paper is as follows. Section 2 briefly reviews the literature. In Section 3, the Johansen methodology is explained. The application of this methodology to the five asset classes is presented in Section 4 along with the results. The last section contains the conclusion and suggestions for further research.

Literature Review

It is already known that some financial returns are predictable by their own history, as well as by the history of other series. In real estate, there is a large body of literature establishing the predictive ability of common stocks and securitized real estate with respect to unsecuritized real estate (See Gyourko and Keim 1992; Myer and Webb 1993; and Barkham and Geltner 1995). However, important relationships can be missed if a study begins with returns. To alleviate this problem, the possibility of cointegrating relationships among wealth indices should be considered. If cointegration is present, a more reliable predictive model for returns can be specified. The long-term relationships have significance for the non-speculative, long-term oriented investor. As recently discussed by Alexander and Johnson (1994), according to modern portfolio theory (MPT), if the correlation between assets is low, there is potential for diversification. However, the existence of a long-term relationship diminishes the opportunity for risk diversification in the long-run. This study considers a more sophisticated problem of the same type, asset substitutability, but in a long-run sense. A long-term relationship might indicate the possibility of substituting one asset class with a combination of other assets. Also, according to the theory of cointegration, it is possible to form a short-run vector error correction model (VECM) where the short-run correction is precisely the error of the long-run relationship. Such a model can be used for forecasting and to assess the predictive power of some asset classes with respect to others. Therefore, the cointegrating relationship is too important to be ignored in studies of Granger causality, in non-structural (VAR) manipulations, and in prediction models.

Methodology

Individual economic series may be non-stationary over time. Linear regression among such series can produce spurious results. The standard practice of taking the first difference, or detrending the series, can lead to misspecifications (See McCallum 1993). However, even if series are non-stationary, they could be integrated (Engle and Granger 1987), and therefore, when considered in a system, it is possible for one or more combinations of the series to be stationary. In such cases, the series are said to be cointegrated. This is in contrast with correlation, which is a static term and does not take into account the ordering in time of the series. Cointegration takes into account the dynamic links among time series. Cointegrated series are tied together in the long-run, but in the short-run deviations from the long-run relationship are possible, provided that these deviations are stationary. Engle and Granger (1987) proved that an error correction term based on the long-run relationship, given a movement away from equilibrium in one period, will adjust a proportion of the disequilibrium in the next period. Further, Engle and Yoo (1987) use simulated data to illustrate possible improvements in long-term forecasting accuracy using error correction models versus unrestricted VAR models, when the variables are known to be cointegrated by design. While the methodology introduced by Engle and Granger (1987) is extremely valuable, it does not reliably deal with the long-run relationship when three or more series are cointegrated.

Johansen (1988) presents a formal exposition of the multicointegration procedure. Since the method is widely used in empirical research [1], only a brief intuitive exposition is given here. The method starts with a nonstructural VAR model. The procedure of correction leads to a vector error correction model (VECM) which effectively is a vector autoregression model(VAR) in differences with r lagged error correction terms included in each equation[2]. The procedure considers two sets of variables: The first set of variables consists of differences that are stationary by definition, whereas the second set consists of lagged levels that are integrated. The Johansen procedure consists of finding the combinations of system variables having the highest correlation with the differenced variables. These combinations will be the cointegrating vectors. In order to find these combinations, canonical correlation is used. The eigenvalues may then be used to construct likelihood ratio tests in order to find the combinations that are distinctive (the cointegrating vectors).

First, a test for identifying the order of integration for each series is used. One test that has gained popularity is the Augmented Dickey-Fuller (ADF) test. This test is widely regarded as being the most efficient[3] test from among several tests for integration and is currently the most widely used in practice. The ADF test forms a model like (1) for each economic series. The null hypothesis is that the coefficients are equal to zero. In the cases where the null hypothesis cannot be rejected, the series are deemed to be integrated and the possibility of cointegration arises[4].

$$\Delta z_t = \delta \cdot z_{t-1} \sum_{i=1}^{k} \delta_i \cdot \Delta z_{t-1} + \varepsilon_t \qquad (1)$$

Once the integration has been established, the next step is to consider an unrestricted VAR model like the following:

$$Z_t = \sum_{i=1}^{k} A_i Z_{t-1} + \varepsilon_t \qquad (2)$$

where Z_t contains all n series of the model and ε_t is a vector of random errors.

The VAR model (2) can also be represented in the form (See Johansen and Juselius 1990):

$$\Delta Z_t = \sum_{i=1}^{k-1} \Gamma_i \Delta Z_{t-k} + \Pi Z_{t-k} + \varepsilon_t \qquad (3)$$

where:

$\Gamma_i = -I + A_1 + \ldots + A_i$ (I is a unit matrix),

$\Pi = -I - A_1 \ldots - A_k$

The transformation of a VAR model into a model like equation (3) is called a cointegrating transformation.

The focus of the Johansen procedure is the rank of matrix II of equation (3). Since there are n variables that constitute the vector Z_t, the dimension of II is $n \times n$ and its rank can be, at most, equal to n. It follows from the Granger Representation Theorem (See Engle and Granger 1987, or Johansen 1988) that under some general conditions:

- If the rank of matrix II is equal to n, the vector process Z_t is stationary (That is, all the variables in Z_t are integrated of order zero);

- If the rank of matrix Π is equal to $r<n$, there exists a representation of Π such that: $\Pi = \alpha \cdot \beta'$ where α and β are both $n \times r$ matrices.

Matrix β is called the cointegrating matrix and has the property that $\beta' Z_t \sim I0$, while $Z_t \sim I1$. The straightforward conclusion is that the variables in Z_t are cointegrated, with the cointegrating vectors $\beta_1, \beta_2, \ldots, \beta_r$ being particular columns of the cointegrating matrix β. Hence, in a VAR model which explains n variables there can be, at most, n -1 cointegrating vectors. Matrix α is known as the adjustment matrix (or the feedback matrix). It indicates the speed with which the short-run deviations from the long-run equilibrium are corrected.

For empirical analysis, the essential problems are in the determination of r (ie.- in identifying the number of cointegrating vectors and in estimating the cointegrating matrix β). Using the canonical correlation procedure, it is possible to identify λ_i, the eigenvalues and compute the LR statistic (also known as the trace test).

$$LR_{trace} = -T \cdot \sum_{i=r+1}^{n} ln(1-\lambda_i) \tag{4}$$

The null hypothesis is that there are (at most) r cointegrating vectors. The test starts from r = 0, the hypothesis that states there are no cointegrating vectors in a VAR model. If this cannot be rejected, the procedure stops since no confirmation of the existence of cointegrating vectors has been found. If it is rejected, it is possible to sequentially examine the hypothesis that $r \le 1, r \le 2$, etc. If the null hypothesis cannot be rejected for $r \le k$, but it has been rejected for $r \le k$-1, the conclusion is that the number of cointegrating vectors, or the rank of β, is k.

Similarly, the likelihood ratio test statistic for testing the null hypothesis, which states that the number of cointegrating vectors is r, versus the alternative that there are r+1 cointegrating vectors, is given by the following (also known as the maximum eigenvalue test):

$$LR_{\lambda\ max} = -T \ln(1-\lambda_{r+1}) \tag{5}$$

Johansen (1988) has shown that the first r estimated eigenvectors are the maximum likelihood estimates of the columns of β, the cointegrating vectors. The cointegrating vectors can be normalized with the coefficient of any variable and are interpreted as the long-run economic relationship.

Research Design and Data

The methodology described above is applied to five series of wealth indices: U.S. Treasury Bills (T-BILL), long-term bonds (LTB), large capitalization common stock (S&P 500), unsecuritized real estate (NCREIF), and securitized real estate-equity REITs (NAREIT). Data for T-BILL, LTB, and S&P 500 are obtained from SBBI 1996 Yearbook, published by Ibbotson Associates. NCREIF is the index published by the National Council of Real Estate Investment Fiduciaries in cooperation with Frank Russell Company, and NAREIT represents the equity REITs from the 1995 REIT Handbook, published by the National Association of Real Estate Investment Trusts. This study uses 72 quarterly observations from 1978 through 1995. The impetus is two-fold. First, it allows for the determination of long-run relationships among the five series of wealth indices[5]. Second, it provides an improved short-term forecasting model, incorporating the error corrections in a VAR model. By starting with a model specified by indices and applying the error correction methodology, the result is a model specified in returns. Taking the log of each variable and then the difference, a model of continuously compounded returns is obtained which incorporates the error correction term as lagged levels (indices of wealth). As such the model below is tested:

$$r_t^i = \sum_{i=1}^{5}\sum_{j=1}^{4}\delta_j r_{t-j}^i \sum_{\lambda=1}^{k}\alpha_\lambda ECM_{\lambda,t-5} + \sum_{v=1}^{4} + QD_v + \varepsilon_{it} \qquad (6)$$

where i = T-BILL, LTB, S&P 500, NCREIF, NAREIT; QD are dummy variables for each quarter; and ECMs are the error correction terms. The choice of four lags was made using the AIC and SC selection criteria. While many other lags were tried, the intuitive choice, given the appraisal induced seasonality for NCREIF, proved to be the best. Seasonal dummy variables were included because some of the variables, particularly NCREIF, showed strong seasonality. Because the model includes all four dummies, it is estimated as if a constant is in the system even though the constant is not explicit. Any of the dummy variables can have the role of the constant. This specification is consistent with the ADF test that consistently rejects the null hypothesis of no constant.

Results

The ADF tests[6] indicate that each of the five series are integrated of order one. The Johansen (1988) procedure was applied to the full data set (72 observations) and the hypothesis of two cointegrating vectors could not be rejected. A VAR model and the corresponding VECM models were then compared for the full data set. Results show that the VECM model is a better fit than an unrestricted VAR model[7]. Since the full model has a reasonable goodness of fit, especially for NCREIF and T- BILL, a prediction model was constructed using 64 quarterly observations with 8 observations (two years) as a holdout sample. The same Johansen (1988) procedure was then applied to the prediction model. Table I presents the results. Clearly, the hypothesis of two cointegrated vectors cannot be rejected through a combined test for the Trace and the Maximum Eigenvalue. As such, the first two columns in Panel B are the cointegrated vectors representing the long-term relationships and the error correction terms at the same time[8]. Panel C presents the feedback matrix where only the first columns are of importance. Johansen (1988) attaches more importance to the first vector because it corresponds to the highest eigenvalue. The second vector does not contain any new information about the long-term relationships. However, since a linear combination of the cointegrating vectors is also a cointegrating vector, a new cointegrating vector that excludes one of the series can be obtained. The first eigenvector normalized for NCREIF leads to the following long-term relationship:

NCREIF=2.296(TBILL)+0.477(S&P500)- (1.490) LTB-0.306(NAREIT) (7)

It is useful to eliminate one of the variables between the first two eigenvectors in order to discover more relationships. For instance, the S&P 500 which in previous research has consistently been shown to have little or no correlation with NCREIF, can be eliminated. The relationship is then:

NCREIF=2.425(TBILL)-0.945(LTB)-0.118(NAREIT) (8)

These relationships are void of economic content until a meaningful equilibrium model that includes the relationships among the group of assets is delineated. This endeavor is well beyond the scope of this study. However, such long-run relationships can bring new information to the long-term investment problem. Both (7) and (8) show that the index of unsecuritized real estate is positively linked to Treasury Bills (T-BILL), and negatively related to long-term bonds (LTB) and securitized real estate (NAREIT).

The portfolio allocation problem involving Treasury Bills, long-term bonds, common stock, and real estate for risk diversification has received considerable attention. Generally, the extant literature considers only the standard MPT problem based on correlations. That is, whenever the inter-asset correlation is low, there is an opportunity for risk diversification. But, if assets are low to moderately correlated and also cointegrated, a new situation emerges. First, it may be possible to hedge some asset classes against others, particularly if their returns move in opposite directions over time. Second, if the goal is to diversify, not all asset classes are necessary. It appears that the system of five asset series is governed by three common (non-stationary) factors since two cointegrating vectors are present (See Stock and Watson 1988).

It was noted by Granger (1988) that cointegration implies "Granger causality" and thus there is potential for a partial forecast of a variable by its past history and the history of the other variables present in the cointegrated system. At first this would seem to be a violation of the weak-form of the efficient market hypothesis. However, the results from this study corroborates those of Mei and Lee (1994), and indicate that the prediction does not necessarily violate market efficiency, but rather that time varying risk premia might be present.

The second part of the results shows how cointegration can help in predicting short-run movements. Table II presents the results for the vector error correction model (VECM), while Table III presents the results for the unrestricted VAR model. Each column gives the β coefficient and their respective probability (p-values). The dependent variable is indicated in the heading and the independent variables are in the first column. The model is estimated using a heteroskedasticity and autocorrelation consistent matrix with order four. The relevant goodness of fit statistics show that the model is doing a very good job for NCREIF and T-BILL, has moderate success with LTB, and is unsuccessful in the case of S&P 500 and NAREIT. Further analysis shows that for NCREIF, LTB, and NAREIT, the VECM model improves the goodness of fit. For T-BILLs and S&P 500 there is little or no change from one model to the next. The results are not surprising, given the significance of the error correction term. It is clear that in the case of T-BILL and S&P 500, the term is not significant, therefore the correction will not adjust the series if they diverge from each other in the short-run. In the case of T-BILL, the correction is not necessary because of the very good fit for the VAR model. In the case of the S&P 500, it is not useful since none of the models provide a good fit. For the asset classes where the error correction plays an important role, it is interesting to note that assets believed to help predict NCREIF (for instance), especially NAREIT, lose their significance when the model incorporates the error correction model. Also, NAREIT is influenced now by lags of NCREIF (as well as other lagged variables). These results stand in contradiction with the existing literature[9] and indicates that ignoring cointegration can lead to false relationships.

In order to explore the forecast capability of the constructed models, eight points are used as a holdout sample. A VAR and VECM forecast are developed and compared with the usual statistics: R square between predicted and actual, sum of square errors, and Theil's U inequality coefficient. Table 4 shows the prediction power of each model for all asset classes. The three statistics presented show that VECM is a superior forcasting model only for the unsecuritized real estate (NCREIF) returns. In the case of T-BILL, there is no difference between the models. This is clear from the non-significant error term in the VECM model. For the other asset classes, the prediction capability of both models is rather poor as indicated by the low goodness of fit. Apparently these last three indices, LTB, S&P 500, and NAREIT, are not amenable to forecasting.

If long-term risk diversification is the goal, from the five asset classes it appears that only three should be included in an efficient frontier. The remaining two assets are long-term linear combinations of the other three. On the other hand, if timing strategies is the investment goal then attention should be given to the two assets that can be forecast, NCREIF and T-BILLs. The results from this study reinforce the initial idea that the long-run disturbance terms improve short-term prediction by correcting the drift from equilibrium.

Conclusions

This study analyses the long-term relationships among five indices of wealth: large capitalization stocks, Treasury Bills, securitized and unsecuritized real estate, and long-term corporate bonds. Based on the observation that most economic time series are integrated, this study discusses the impact that such a relationship could have for investment decisions. Two cointegrating relationships are discovered among the five asset classes indicating that the asset classes move together over time. The results indicate that a reasonably accurate forecast is possible for the unsecuritized real estate (NCREIF). Also, for unsecuritized real estate, an improvement in the performance is realized when the VECM is used. The inclusion of the correction term changes significantly some of the causalities among real estate returns previously established in the literature. Evidence of securitized real estate having predictive power for the unsecuritized real estate, as documented initially in Gyourko and Keim (1992), is not supported. On the contrary, evidence that securitized real estate is weakly influenced by lagged unsecuritized real estate, along with other asset classes is obtained. At first glance, the evidence of prediction can be taken as evidence of weak-form of the efficient market hypothesis. However, the existence of time varying risk premia cannot be dismissed. Therefore more tests are necessary in order to conclude that markets are inefficient.

If the goal of investing is risk diversification, a long-term relationship could make some asset classes unnecessary. If trading strategies are the goal, then one can take advantage of the improved predictive ability of a VECM model over a simple VAR model. Further research should consider these ideas in more detail. On the risk side, efficient frontiers for all five classes should be compared with efficient frontiers obtained by eliminating the assets indicated by the long-run relationships and replacing them with the relationship itself. The comparison should, however, consider a long-term horizon because the posited relationship is a long-term one. On the investing strategy side, trading rules based on VECM forecasts should be compared with those obtained using VAR or naive models. The cointegration relationships shown in this study should have a practical impact concerning the way asset allocation is performed and trading strategies implemented.

References

Alexander C. and A. Johnson, 1994, Dynamic links, *Risk*, February, pp 56-59.

Barkham R. and D. Geltner, 1995, Price discovery in American and British property markets, *Real Estate Economics*, 23 pp. 21-44.

Engle R. and C. Granger, 1987, Co-integration and error correction: representation, estimation, and testing, *Econometrica*, 55 pp. 251-276.

Engle R. and B. Yoo, 1987, Forecasting and testing in co-integrated systems, *Journal of Econometrics*, 35 pp. 143-159.

Geltner D. and J. Mei, 1995, The present value model with time-varying discount rates: implications for commercial property valuation and investment decisions, *Journal of Real Estate Finance and Economics*, 11 pp. 119-135.

Granger C., 1988, Some recent developments on the concept of causality, *Journal of Econometrics* ,39 pp. 199-211.

Gyourko J. and D. Keim, 1992, What does the stock market tell us about real estate returns, *AREUEA Journal*, 20 pp. 457-485.

Hargreaves C.P., editor, 1994, *Nonsationary time series analysis and cointegration*, Oxford University Press.

Johansen S., 1988, Statistical analysis of the cointegrating vectors , *Journal of Economic Dynamics and Control*, 12 pp. 231-254.

Johansen S. and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration — with applications to the demand for money, *Oxford Bulletin of Economics and Statistics*, pp. 169-210.

McCallum B.T., 1993, Unit roots in macroeconomic time series: some critical issues, NBER Working Paper No. 4368

Mei J. and A. Lee, 1994, Is there a real estate factor premium?, *Journal of Real Estate Finance and Economics*, 9 pp. 113-126.

Myer F. and J. Webb, 1993, Return properties of equity REITs, common stocks, and commercial real estate: a comparison, *Journal of Real Estate Research*, 8 pp. 87-106.

Stock J. and M. Watson, 1988, Testing for common trends, *Journal of the American Statistical Association*, 83 pp. 1097-1107.

TABLE I
Cointegration Tests and Results

PANEL A: Tests and Critical Values

r	Max λ	Critical Values		Trace	Critical Values	
		95%	97.50%		95%	97.50%
4	0.498	8.18	9.72	0.498	8.18	9.72
3	9.713	14.9	17.07	10.211	17.95	20.08
2	*16.232	21.07	22.89	*26.444	31.25	34.48
1	45.882	27.14	29.16	72.325	48.28	51.54
0	62.469	33.32	35.8	134.794	70.6	74.04

PANEL B: Vectors of the Cointegrating Matrix

β_1#	β_2#	β_3	β_4	β_5
18.9893	-4.6908	17.6059	0.6031	24.8203
-49.3088	-4.2858	-21.5594	-37.4113	-14.5736
-9.0729	-43.8134	0.0314	-3.6127	-6.3722
28.3048	45.5840	9.5142	8.3816	-8.2779
5.8171	16.7010	-9.5272	14.9858	0.0292

PANEL C: Vectors of the Feedback Matrix

α_1#	α_2#	α_3	α_4	α_5
-0.0023	-0.0012	-0.0014	0.0027	0.0002
0.0002	-0.0007	0.0001	-0.0001	0.0001
0.0014	0.0319	0.0177	0.0016	0.0027
-0.0117	0.0150	-0.0054	-0.0048	0.0003
-0.0312	0.0234	0.0134	0.0007	0.0010

Panel A shows the obtained and critical values for the maximum eigenvalue and trace test. Panel B shows the cointegrating vestors; the significant eigenvectors are marked with #. Panel C shows the feedback matrix; the relevant adjustment coefficients are marked with #. The null hypothesis for the trace test is: the no. of cointegrating vectors is less than or equal to r, with the alternative of r+1. The null hypothesis for the maximum eigenvalue test is: the no. of cointegrating vectors is r, with the alternative of r+1. The * indicates when the null hypothesis is not rejected in the two tests considered.

TABLE II Results for the Vector Error Correction Model (VECM)					
	T-Bill	LTB	S&P 500	NCREIF	NAREIT
NCR[-1]	0.0392	0.0650	-0.3625	0.1020	-1.0625
	***(0.0010)	(0.7990)	(0.6230)	**(0.0480)	*(0.1070)
TBILL[-1]	1.0220	-0.0397	2.5811	0.3423	3.2973
	***(0.0000)	(0.9850)	(0.6870)	**(0.5000)	(0.4580)
SP[-1]	-0.0032	0.1367	-0.1516	0.0050	0.3156
	(0.5520)	(0.1220)	(0.6250)	(0.8200)	(0.1520)
LBGC[-1]	-0.0382	0.0527	0.5949	-0.1031	0.5288
	***(0.0000)	(0.7190)	***(0.0080)	**(0.0230)	***(0.0080)
NAR[-1]	0.0100	-0.4075	-0.0734	0.0400	-0.5452
	*(0.0960)	***(0.0050)	(0.8170)	(0.1140)	***(0.0110)
NCR[-2]	0.0198	-0.2505	0.4110	0.2022	0.6187
	(0.1650)	(0.4300)	(0.4290)	***(0.0010)	(0.1860)
TBILL[-2]	-0.3256	-1.4408	-10.5540	-0.1432	-11.7470
	*(0.0730)	(0.5960)	(0.1730)	(0.8510)	*(0.0850)
SP[-2]	-0.0004	0.0721	-0.1868	-0.0188	0.1668
	(0.8970)	(0.5770)	(0.4380)	(0.5110)	(0.3840)
LBGC[-2]	0.0081	0.2054	0.7229	-0.0111	0.1688
	(0.4950)	(0.1740)	(0.1640)	(0.7020)	(0.6200)
NAR[-2]	0.0078	-0.4088	-0.1353	0.0301	-0.3787
	*(0.0850)	*(0.0830)	(0.6010)	(0.4250)	*(0.0860)
NCR[-3]	-0.0111	-0.4069	0.2273	0.0978	0.3732
	(0.4210)	(0.1730)	(0.7290)	*(0.1050)	(0.4750)
TBILL[-3]	0.1583	-0.0440	7.0376	-0.2910	5.5945
	(0.4590)	(0.9890)	(0.4320)	(0.6800)	(0.4180)
SP[-3]	-0.0021	-0.0648	-0.3609	0.0437	0.0611
	(0.5820)	(0.4750)	(0.1610)	*(0.0900)	(0.7000)
LBGC[-3]	-0.0059	-0.1369	-0.2563	0.0350	-0.8217
	(0.4390)	(0.3040)	(0.6180)	(0.5330)	**(0.0370)
NAR[-3]	0.0014	0.2247	0.4503	-0.0872	0.0591
	(0.7240)	*(0.0600)	*(0.0680)	(0.1340)	(0.7530)
NCR[-4]	-0.0016	0.7911	1.1800	0.6870	1.5153

	(0.8640)	***(0.0050)	(0.1660)	***(0.0000)	***(0.0140)
TBILL[-4]	0.0890	-1.8046	-1.8195	-1.1415	-10.1190
	(0.5470)	(0.4010)	(0.7770)	**(0.0240)	**(0.0290)
SP[-4]	0.0023	0.1576	-0.1920	0.0110	0.3983
	(0.7410)	(0.1310)	(0.5140)	(0.6070)	**(0.0250)
LBGC[-4]	-0.0128	-0.0427	0.4671	0.0573	-0.0753
	(0.3680)	(0.7900)	(0.2620)	(0.2480)	(0.8030)
NAR[-4]	0.0015	-0.3719	-0.0692	0.0245	-0.7978
	(0.7990)	***(0.0000)	(0.8670)	(0.3570)	***(0.0030)
Q1	0.0006	0.0984	0.0254	0.0205	0.2565
	(0.6990)	***(0.0040)	(0.8140)	**(0.0270)	***(0.0010)
Q2	0.0010	0.1001	0.0615	0.0217	0.3433
	(0.4570)	***(0.0030)	(0.5090)	***(0.0110)	***(0.0000)
Q3	0.0001	0.1258	0.0518	0.0154	0.2957
	(0.9330)	***(0.0020)	(0.6180)	*(0.0920)	***(0.0000)
Q4	0.0005	0.1148	0.0427	0.0193	0.2745
	(0.7500)	***(0.0060)	(0.6380)	**(0.0400)	***(0.0000)
ECM[-5]	0.0002	-0.0117	0.0014	-0.0023	-0.0312
	(0.3010)	***(0.0070)	(0.8900)	**(0.0260)	***(0.0000)
Adj. R2	0.9362	0.2389	-0.2021	0.7201	0.1499
AIC	-12.2780	-6.3144	-4.6627	-8.7964	-5.1847
SC	-11.3980	-5.4341	-3.7824	-7.9161	-4.3044
F	302.9790	2.9580	1.0190	15.8220	2.2220
p-value	0.0000	0.0020	0.4730	0.0000	0.0150

The Table presents the βs of the VECM forcasting model. Each equation is indicated by the appropriate label. The OLS regressions were corrected for heteroskedasticity and autocorrelation with the NEWEY-WEST (1987) method with four lags. The model is estimated without a constant, but includes one since all the quarterly dummies are in the equation. AIC is AKAIKE (1973) information criterion and SC is the Schwartz (1978) criterion for model selection. The significance level of p-values is indicated by asterisks as follows: *** $p \leq 1\%$, ** $1\% < p \leq 5\%$, * $5\% < p \leq 10\%$

TABLE III
Results of the Unrestricted Vector Autoregression Model

	T-BILL	LTB	S&P 500	NCREIF	NAREIT
NCR[-1]	0.0388	0.0928	-0.3659	0.1075	-0.9880
	***(0.0010)	(0.7360)	(0.6150)	*(0.0870)	(0.2700)
TBILL[-1]	1.0069	1.0871	2.4422	0.5650	6.3182
	***(0.0000)	(0.6320)	(0.7060)	(0.2960)	(0.2360)
SP[-1]	-0.0009	-0.0330	-0.1306	-0.0285	-0.1393
	(0.8130)	(0.6870)	(0.5310)	(0.1160)	(0.3670)
LBGC[-1]	-0.0380	0.0388	0.5966	-0.1058	0.4916
	***(0.0000)	(0.7780)	***(0.0080)	**(0.0240)	***(0.0100)
NAR[-1]	0.0070	-0.1854	-0.1008	0.0839	0.0502
	(0.1100)	*(0.0830)	(0.6080)	***(0.0000)	(0.7700)
NCR[-2]	0.0201	-0.2760	0.4142	0.1971	0.5504
	(0.1650)	(0.3020)	(0.4330)	***(0.0020)	(0.2890)
TBILL[-2]	-0.3333	-0.8706	-10.6240	-0.0305	-10.2180
	*(0.0730)	(0.7830)	(0.1600)	(0.9680)	(0.1950)
SP[-2]	0.0015	-0.0734	-0.1689	-0.0476	-0.2231
	(0.6220)	(0.4890)	(0.3890)	**(0.0530)	(0.2230)
LBGC[-2]	0.0065	0.3286	0.7077	0.0133	0.4992
	(0.5590)	**(0.0510)	(0.1470)	(0.7120)	*(0.0770)
NAR[-2]	0.0054	-0.2318	-0.1571	0.0650	0.0956
	(0.1990)	(0.2340)	(0.4980)	**(0.0310)	(0.6240)
NCR[-3]	-0.0083	-0.6198	0.2536	0.0557	-0.1977
	(0.5030)	**(0.0240)	(0.6570)	(0.4380)	(0.6990)
TBILL[-3]	0.1684	-0.8025	7.1310	-0.4409	3.5614
	(0.4410)	(0.8190)	(0.4230)	(0.5680)	(0.6350)
SP[-3]	-0.0007	-0.1708	-0.3478	0.0228	-0.2232
	(0.8510)	*(0.0930)	*(0.0580)	(0.2910)	**(0.0540)
LBGC[-3]	-0.0081	0.0238	-0.2761	0.0668	-0.3910
	(0.2280)	(0.8590)	(0.5450)	(0.3130)	(0.3270)
NAR[-3]	0.0000	0.3285	0.4375	-0.0667	0.3376

	(0.9930)	**(0.0190)	**(0.0270)	(0.2250)	(0.1130)
NCR[-4]	0.0005	0.6364	1.1990	0.6564	1.1005
	(0.9590)	**(0.0180)	(0.1380)	***(0.0000)	*(0.0730)
TBILL[-4]	0.0442	1.5354	-2.2312	-0.4815	-1.1652
	(0.7740)	(0.4080)	(0.6900)	(0.2900)	(0.7870)
SP[-4]	0.0044	-0.0029	-0.1722	-0.0207	-0.0318
	(0.4470)	(0.9770)	(0.4000)	(0.4640)	(0.8440)
LBGC[-4]	-0.0144	0.0736	0.4528	0.0803	0.2363
	(0.3180)	(0.6440)	(0.2490)	(0.1220)	(0.4670)
NAR[-4]	-0.0013	-0.1621	-0.0951	0.0659	-0.2352
	(0.7970)	*(0.0830)	(0.7480)	(0.1120)	(0.1870)
Q1	0.0017	0.0116	0.0361	0.0033	0.0236
	*(0.0630)	(0.5330)	(0.4320)	(0.5790)	(0.5330)
Q2	0.0022	0.0079	0.0728	0.0035	0.0962
	***(0.0060)	(0.6320)	**(0.0500)	(0.5190)	***(0.0130)
Q3	0.0014	0.0309	0.0635	-0.0034	0.0413
	(0.2010)	(0.1400)	*(0.0960)	(0.5550)	(0.2600)
Q4	0.0017	0.0250	0.0538	0.0016	0.0337
	*(0.0590)	(0.2220)	(0.1640)	(0.7510)	(0.4180)
Adj. R2	0.9372	0.1331	-0.1683	0.7058	-0.1617
AIC	-12.3000	-6.1869	-4.6961	-8.7516	-4.8773
SC	-11.4550	-5.3418	-3.8510	-7.9065	-4.0322
F	320.7300	2.4810	1.0910	15.5720	1.2720
p-value	0.0000	0.0070	0.4000	0.0000	0.2530

The Table presents the βs of the unrestricted VAR forcasting model. Each equation is indicated by the appropriate label. The OLS regressions were corrected for heteroskedasticity and autocorrelation with the NEWEY-WEST (1987) method with four lags. The model is estimated without a constant, but includes one since all the quarterly dummies are in the equation. AIC is AKAIKE (1973) information criterion and SC is the Schwartz (1978) criterion for model selection. The significance level of p-values is indicated by asterisks as follows:

*** $p \leq 1\%$, ** $<1\%\ p \leq 5\%$, * $<5\%\ p \leq 10\%$

TABLE IV
Statistical Significance of the Forecasting models

	T-BILL		LTB		S&P 500		NCREIF		NAREIT	
	VECM	VAR	VECM	VAR	VECM	VAR	VECM	VAR	VECM	VAR
SSE	.0001	.0001	.0309	.0152	.0222	.0228	.0015	.0025	.1291	.0214
Theil U	1.022	.711	2.218	1.48	.1.082	1.067	2.083	2.794	2.255	.765
R^2	.86	.90	.06	.00	.00	.00	.41	.29	.53	.02

SEE is the sum of squared errors between actual values and the forecast. R^2 is com puted between actual and predicted values. Theil U is the Theil's inequality coeffi cient, a measure of forecast accuracy defined by the following formula:

$$U = \frac{\sqrt{1/T\sum(Y_t^f - Y_t^a)^2}}{\sqrt{1/T\sum(Y_t^f)^2} + \sqrt{1/T\sum(Y_t^a)^2}}$$

where Y^f is the forecast and Y^a is the actual value, and T is the total number of predictions.

Endnotes

1. For a comprehensive theoretical and empirical collection of studies discussing this methodology and its limitations, see Hargreaves (1994).

2. The following classic notations will be used throughout. The operator Δ in the expression of the type ΔZ_t will denote first differencing of all the variables in a vector of variables Z_t; the notation $Z_t \sim I1$ means that all the variables which constitute the vector of variables Z_t are integrated of order one and therefore their first differences are stationary. Notation $Z_t \sim CI1,1$ with cointegrating vector β means that the linear combination $\beta' Z_t \sim I0$. In the case where more than one cointegrating vector exists, $\beta_1, \beta_2, ..., \beta_l$ constitute a matrix β such that $\beta' Z_t \sim I0$.

3. Other tests are Dickey-Fuller and Pillips-Peron. All these tests were used in this study and lead to the same conclusion.

4. z_i represents a vector (a time series) from the matrix Z of time series.

5. The importance of the wealth indices for economic decisions cannot be stressed enough. Wealth effects on consumption, as well as Tobin's q theory of investment, represent important monetary transmission mechanisms.

6. Results are not reported for brevity, but are available from the authors by request. The failure to reject the null hypothesis of integration for time series is confirmed in the literature.

7. This study focuses on the prediction model with a holdout sample of eight points. The results for the full set (72 points and for the 64 points used in the prediction model) are very similar. All results are available upon request from the authors.

8. The error correction term is obtained by multiplying each coefficient by its respective variable and by algebraically summing the products. It is desirable for the mean of the error correction to be zero. Therefore, the mean of each obtained error correction was subtracted from the error correction term.

9. Gyourko and Keim(1992), Barkham and Geltner (1995) and Geltner and Mei (1995) are some of the studies indicating that unsecuritized property returns are forecast by securitized real estate.