A study on the volatility forecast of the US housing market in the 2008 crisis

Kui-Wai Li

2011

Applied Financial Economics, Vol. 22, No. 22, November 2012, pp. 1869-1880.

A Study on the Volatility Forecast of the U.S. Housing Market in the 2008 Crisis

Kui-Wai Li
City University of Hong Kong

Abstract

This article provides the in-sample estimation and evaluates the out-of-sample conditional mean and volatility forecast performance of the conventional GARCH, APARCH and the benchmark Riskmetrics model on the U.S. real estate finance data for the pre-crisis and post-crisis periods in 2008. The empirical results show that the Riskmetrics model performed satisfactorily in the in-sample estimation but poorly in the out-of-sample forecast. For the post-crisis out-of-sample forecasts, all models naturally performed poorly in conditional mean and volatility forecast.

JEL Classification: C53, G17, L85, R21

Keywords: financial crisis, volatility forecast, U.S. real estate finance

———

Corresponding author: Kui-Wai Li, Department of Economics and Finance, City University of Hong Kong. Tel.: 852-34428805; fax: 852-34420195; E-mail address: efkwli@cityu.edu.hk

Acknowledgement: The author would like to thank the Editor, Mark Taylor, and two anonymous referees for their invaluable comments on the earlier draft of the paper. Research funding from the City University of Hong Kong (Strategic Research Grant numbers 7002523 and 7008129) is gratefully acknowledged. The two research assistants, Siyang Ye and Douglas K. T. Wong, have provided useful research inputs. The author is responsible for mistakes found in the paper.

I Introduction

Most analysts would agree that the origin of the September 2008 financial crisis in U.S. began in March 2007 with the collapse of the subprime mortgage industry that eventually led to a worldwide credit crunch as banks and hedge funds invested heavily in subprime mortgage backed securities. The crisis turned imminent when the U.S. Federal Reserve took over the two largest mortgage based security companies, and further deteriorated following the collapse of Lehman Brothers. The subsequent financial meltdown had spread to other international financial centers.

While a number of recent studies on the causes of the 2008 U.S. financial crisis have concentrated on both financial and monetary fundamentals (Taylor and Williams, 2008; Taylor, 2009; Schwartz, 2009; Financial Services Authority, 2009; French *et al*., 2010; Wong and Li, 2010), the crash in the subprime mortgage industry in 2007 could be the result of cumulative economic events since the mid-1990s, including the prolonged low interest rate regime, the rapid recovery in the U.S. housing markets after the burst of the dot-com bubble in 2001 and the continuous increase in unsecured sub-prime lending to unqualified home purchasers. These incidences have led to an increase in the volume of collateralized debt obligations (CDOs) backed by asset and sub-prime mortgage securities. Although the U.S. housing price has accelerated significantly between 2003 and 2006, after the collapse of the subprime mortgage market in early 2007, the large number of mortgage delinquencies and defaults has reversed both the housing price and the value of mortgage-backed securities. The sub-prime mortgage crisis has drawn renewed attention on how banks measure the risk of assets and the extent of accuracy in volatility forecast.

A commonly used approach in measuring investment risk is based on the historical variability of assets return. Markowitz (1952) first used assets return volatility as a measurement of risk. However, Wheaton *et al*. (1999) argued that real estate risk measurement should not be based solely on the historical data, because most real estate assets are still privately owned and do not produce an efficient asset pricing. Under inefficient asset pricing, positive shocks could probably set off asset price fluctuations that would easily be predicted. Wheaton *et al*. (1999) showed that historic variability can be decomposed into predictable and non-predictable components. The predictable components appeared when the series exhibited autoregressive characteristics or other significant patterns. The non-predictable components involved future uncertainty or the asset returns yielded a random walk series. If the asset returns exhibited random walk, the future of the asset returns could not be forecasted.

Undoubtedly, assets returns volatility is still the main concern of investors, financial institutions and regulatory authorities. In empirical financial literature, the most frequent instrument used to measure risk is the Autoregressive Conditional Heteroscedastic approach (ARCH) (Engle, 1982) that allows for the conditional means and variance to change over time. The risk management group of J. P. Morgan (1996) released a technical model, called RiskMetrics, to measure assets returns volatility and volatility forecast. The RiskMetrics model is based on the integrated and generalized ARCH (IGARCH) model of Engle and Bollerslev (1986) with fixed ARCH and GARCH coefficients. In practice, most of the major financial institutions have adopted the RiskMetrics model to manage market risk. The accuracy of the model, therefore, has become the major concern, especially for the out-of-sample forecast. Since the banks would ultimately bear the risk from loan defaults, it has been criticized that banks have under-estimated the potential risk in the sub-prime mortgages.

Recent developments in forecasting analysis can be used to evaluate the accuracy of the various forecast models (West, 1996, 2001, 2006). In this paper, the in-sample estimation of the real estates related financial data series are compared with the out-of-sample conditional mean and volatility forecast performance of the conventional GARCH model, the Asymmetric Power ARCH (APARCH) model and the benchmark Riskmetrics model for the two pre-crisis and post-crisis periods. Section II shows the methodology, while Section III provides the data description. Section IV presents the in-sample estimation and out-of-sample forecast of various models. Section V concludes the paper.

II Methodology

In-sample analysis

The Autoregressive Conditional Heteroscedastic (ARCH) model permits the conditional means and variance to change over time (Engle, 1982). The ARCH model in the in-sample analysis can be given as:

$$y_t \,|\Omega_{t-1} \sim (x_t\beta, h_t) \quad , \tag{1}$$

$$h_t = w_0 + \alpha_1 u_{t-1}^2 \, , \tag{2}$$

where $x_t \beta$ is the mean of y_t, which is a linear combination of lagged variables included in the information set (Ω_{t-1}) with a vector (β) of unknown parameters. Based on past forecast errors, the underlying forecast variance (h_t) may change over time, thereby keeping the unconditional variance constant.

The generalized ARCH (GARCH) model that incorporated the problem of parsimony (Bollerslev, 1986) allows a longer memory and a more flexible structure. The variance equation of the GARCH (p, q) process can be defined as:

$$h_t = w_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \, , \tag{3}$$

where $w_0 > 0$, $\alpha_i \geq 0$ and $\beta_i \geq 0$ ($\forall i$). The GARCH (p,q) process permits an autoregressive moving average component in the heteroscedastic variance.

An alternative to the GARCH-type model is the Asymmetric Power ARCH (APARCH) model (Ding *et al*., 1993) that extends Equation (3) into the following:

$$\sigma_t^{\delta} = w + \sum_{i=1}^{q} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\delta} , \tag{4}$$

where $\delta > 0$ and $-1 < \gamma_i < 1$. The parameter δ plays the role of a Box-Cox transformation of σ_t, while γ_i represents the asymmetric responses that a negative shock to a financial time series is likely to cause higher volatility than a positive shock of the same magnitude (Black, 1976; Christie, 1982; French *et al*., 1987; Nelson, 1991; Schwert, 1990; Engle and Ng, 1993). Taking into account the asymmetric response, the APARCH model allows for the flexibility of a varying exponent. It nests the GJR model (Glosten *et al*., 1993) when $\delta = 2$, and the GARCH model of Taylor (1986) and Schwert (1990) when $\delta = 1$, and $\gamma_i = 0\ (i = 1, \ldots, p)$.

Due to uncertainty in financial markets, the management of financial risk is common in most financial institutions. The risk management group of J. P. Morgan (1996) has proposed a market risk management methodology known as RiskMetrics to manage the potential risk in financial markets. The RiskMetrics model is defined as:

$$\sigma_t^2 = \omega + (1 - \lambda)\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2, \qquad (5)$$

where ω is equal to zero and λ is generally set to 0.094 in practice. Equation (5) is a basic conditional variance model of the RiskMetrics, and the variance equation is modeled as a linear combination of lagged squared residuals and lagged conditional variances. The in-sample estimation of the GARCH, APARCH and Riskmetrics models on the real estates related data series will be compared.

Out-of-sample volatility forecasting

The accuracy in forecasting is important for investors, financial institutions and regulatory authority to measure the potential risk of their asset portfolios. In order to compare the forecasting performance of alternative models, an out-of-sample forecast by the moving window procedure that began with the estimation of each individual model using in-sample period data was used to predict the one-step-ahead (month) volatility forecasts, and the in-sample estimation period was shifted forward by one period for estimation and prediction. This process is repeated *N* times until the last observation of the forecasting period. The predicted one-month-ahead volatility is then compared with the realized volatility and all the estimated results are recorded for the models comparison using the statistical tests.

Though most studies have used the square return as a proxy for volatility (Brailsford and Faff, 1996; Brooks and Persands, 2002; Sadorsky 2006), Andersen and Bollerslev (1998) used the integrated volatility as a proxy for realized volatility. By using this method, the realized volatility can provide a consistent non-parametric estimate of the price variability that has transpired over a given discrete interval. In addition, similar to the other literatures on conditional volatility forecasting, the monthly forecast errors generated from each model are compared by using the following two statistic tests to evaluate and compare the forecast errors between models.

a) Traditional loss functions

The general symmetric loss functions includes the mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE), which are defined, respectively, as:

$$MAE = \frac{1}{n}\sum_{t=T+1}^{T+n}\left|\sigma_{t,f}^{2} - \sigma_{t}^{2}\right|, \tag{6}$$

$$MAPE = \frac{1}{n}\sum_{t=T+1}^{T+n}\left|\left(\sigma_{t,f}^{2} - \sigma_{t}^{2}\right)/\sigma_{t}^{2}\right|, \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=T+1}^{T+n}\left(\sigma_{t,f}^{2} - \sigma_{t}^{2}\right)^{2}}, \tag{8}$$

where n represents the number of forecast. $\sigma_{t,f}^{2}$ and σ_{T}^{2} represent the one-month-forecast volatility and realized volatility, respectively. However, the symmetric loss function assumed that the investors put the same weight on both the over-prediction and under-prediction of volatility. This, however, is not the case in practice.

b) Asymmetric loss functions:

The mean mixed error statistics that considered under-prediction (MME(U)) more heavily is applied in order to account for the asymmetric properties in the loss function. The MME(U) is defined as:

$$MME(U) = \frac{1}{n}\left[\sum_{t=1}^{O}\left|\sigma_{t,f}^{2} - \sigma_{t}^{2}\right| + \sum_{t=1}^{U}\sqrt{\left|\sigma_{t,f}^{2} - \sigma_{t}^{2}\right|}\right], \tag{9}$$

and the statistic with heavier weight on over-prediction (MME(O)) is defined as:

$$MME(O) = \frac{1}{n}\left[\sum_{t=1}^{O}\sqrt{\left|\sigma_{t,f}^{2} - \sigma_{t}^{2}\right|} + \sum_{t=1}^{U}\left|\sigma_{t,f}^{2} - \sigma_{t}^{2}\right|\right]. \tag{10}$$

III Data Description

Since one of the causes of the 2008 financial crisis was the U.S. housing market, the data used in the empirical analysis include the housing price index (HPI), total home market amount (RHMA) and loan to price ratio (LTP) to reflect the housing market situation in U.S. at that time of the 2008 crisis. Some of the commercial and global investment banks (for example, Bear Stearns, Lehman Brothers) in the 2008 financial crisis were faced with an unprecedented loss due to their large involvement in the subprime and other low-rated mortgage securities. Those risky loans are included in the bank's asset account. The commercial bank assets in the form of consumer loan (CL) should, therefore, be investigated while inter-bank loan (IL) could provide information on those risky mortgages transferred to other banks. The choice of variables is meant to provide good proxies for the U. S. housing market that has caused the financial crisis. The use of loan to price ratio, consumer loan and inter-bank loan can help to trace the flow of the risky mortgages in the analysis/

All the monthly data can be obtained from Data Stream and are expressed in logarithm. For the data on RHMA, LTP, CL and IL, the sample period started from January 1988 to February 2009, while HPI covered the period from January 1991 to February 2009. The RHMA, CL and IL are calculated by deducting the log consumer price index (CPI) so as to account for inflation. All the data are first differenced by using the formula $y_t = P_t - P_{t-1}$.

Table 1 shows the descriptive statistics for HPI, RHMA, LTP, CL and IL. Most of the data series exhibit a non-zero Skewness and a high Excess Kurtosis property, and consequently the Jarque-Bera tests for normality are strong and statistically significant, with the exception of HPI. The standard deviation of RHMA is slightly higher than others due to the relatively higher value. The Box-Pierce test $Q(5)$ for returns cannot reject the null that no serial correlation existed in the case of HPI, RHMA, LTP and CL. The results of Box-Pierce test for squared returns $Q^2(5)$ indicated that a strong presence of an ARCH-structure existed in most of the series, with the exception of CL, and the statistic results of the Augmented Dickey-Fuller (ADF) test indicate that all series are stationary after first differential.

Table 1 Descriptive statistics

	HPI	RHMA	LTP	CL	IL
min	-0.004	-0.216	-0.039	-0.027	-0.135
mean	0.004	0.003	0.000	0.001	0.001
max	0.014	0.197	0.041	0.053	0.182
std.dev	0.003	0.053	0.010	0.011	0.043
Skewness	0.133	-0.214	0.024	0.790	-0.010
Excess Kurtosis	-0.406	1.949	2.325	2.554	1.948
Jarque-Bera	1.868	27.996	38.543	26.645	29.580
	(0.393)	(0.000)†	(0.000)†	(0.000)†	(0.000)†
Q (5)	129.16	27.37	14.73	37.01	4.95
	(0.000)†	(0.000)†	(0.0012)*	(0.000)†	(0.422)
Q^2 (5)	120.21	31.09	12.88	9.37	15.94
	(0.000)†	(0.000)†	(0.025)*	(0.095)	(0.007)†
ADF	-6.933†	-4.996†	-4.809†	-2.917†	-5.512†

Note: $Q(5)$ and $Q^2(5)$ represent the Box-Pierce test statistics at lag 5 for returns and required returns, respectively. † and * represent statistical significance at 5% and 10%, respectively.

IV Empirical Results

<u>In-sample analysis</u>

One can start with the in-sample estimation for each series in order to compare the performance between the models. Since the first burst of the sub-prime mortgage crisis occurred in October 2007, the end date prior to the crisis is selected. The in-sample period started from February 1988 to June 2007 for RHMA, LTP, CL and IL, and the period from February 1991 to June 2007 for HPI.

Table 2 reports the estimated parameters of the univariate APARCH, GARCH and Riskmetrics with skewed student distribution. The three parameters of w, α and β are the GARCH parameters from Equation (3), and ϕ_i are the coefficients of the AR process. Since the first lagged value is statistically significant, a total of four series can be specified as AR (1) process in the conditional mean equation, with the exception of IL. As for the conditional variance equation, the weights of α_I and β_I satisfy the non-negativity constraint and the $\alpha_I + \beta_I < 1$ restriction in all models. It is apparent that most of the constant terms are statistically insignificant at 5 percent level, while the coefficients of lagged squared residual and coefficients of lagged variance terms are highly significant statistically and the sum of alpha and beta is close to unity, indicating that the persistence of the conditional variance in all series is high.

In the APARCH estimations, the estimated asymmetric parameter γ_1 is statistically insignificant in all series except HPI. This suggests that there is no asymmetric response to positive and negative shocks in RHMA, LTP, CL and IL. The hypotheses of the parameter δ=1 (conditional standard deviation) and δ=2 (conditional variance) cannot be rejected in the cases of HPI, RHMA and LTP at 5% significant level. This indicates that the conventional GARCH specification may be more appropriate than the APARCH model. In the case of CL and IL, the parameter δ is statistically significant. This result supports the use of a model that allows the power term to be estimated, the APARCH model is more suitable in all cases.

The statistical results of student distribution (*Df*) are highly significant in most series, and the normality test gives the identical results that most series do not follow normal distribution, with the exception of HPI. The Ljung-Box test is used to detect the misspecification in the conditional mean and the variance equation. The results of the portmanteau test on standardized residuals *LB*(8) are mostly insignificant statistically at 5% significant level, with the exception of CL, indicating that the serial correlations in conditional mean have successfully been eliminated by the AR process. Similarly, no serial correlation in variance equation is detected as the results of the portmanteau test on squared standardized residuals $LB^2(8)$ are all statistically insignificant, with the exception of CL.

Table 2 Model estimated results and diagnostic tests

	Skewed APARCH					Skewed GARCH					Skewed Riskmetrics				
	HPI	RHMA	LTP	CL	IL	HPI	RHMA	LTP	CL	IL	HPI	RHMA	LTP	CL	IL
GARCH parameters															
ϕ_0	0.004	0.002	0.0000	0.001	0.002	0.004	0.002	0	0.001	0.001	0.004	0.003	0	0.000	0.001
	(0.001)†	(-0.002)	(0.000)	(0.001)	(0.003)	(0.001)†	(0.002)	0	(0.001)	(0.003)	(0.001)†	(0.002)	(0.001)	(0.001)	(0.003)
ϕ_1	0.622	-0.351	-0.218	0.395	0.018	0.618	-0.33	-0.226	0.400	0.021	0.662	-0.314	-0.21	0.406	0.015
	(0.05)†	(0.073)†	(0.08)†	(0.056)†	(0.071)	(0.058)†	(0.068)†	(0.066)†	(0.059)†	(0.074)	(0.055)†	(0.078)†	(0.066)†	(0.062)†	(0.078)
w_1	14.462	58.408	18.419	6.831	15.216	3.139	3.862	8.713	0.010	2.383					
	(80.66)	(114.5)	(117.39)	(70.11)	(89.02)	(2.275)	(1.552)†	(4.812)*	(0.017)	(4.776)					
α_1	0.06	0.263	0.175	0.027	0.088	0.076	0.234	0.17	0.026	0.089	0.06	0.06	0.06	0.06	0.06
	(0.041)	(0.065)†	(0.078)†	(0.015)*	(0.092)	(0.094)	(0.085)†	(0.078)†	(0.015)*	(0.096)					
β_1	0.823	0.633	0.749	0.959	0.813	0.523	0.648	0.749	0.97	0.786	0.94	0.94	0.94	0.94	0.94
	(0.093)	(0.075)†	(0.082)†	(0.023)	(0.295)†	(0.333)	(0.068)†	(0.07)†	(0.028)†	(0.34)†					
γ	0.866	0.102	0.106	0.071	-0.237										
	(0.391)†	(0.301)	(0.234)	(0.154)	(0.264)										
δ	1.522	1.137	1.849	2.043	1.999										
	(0.956)	(0.61)†	(1.35)	(0.076)†	(0.043)†										
θ	0.131	-0.102	-0.047	0.207	-0.048	0.138	-0.097	-0.033	0.249	-0.071	0.168	-0.037	-0.02	0.185	-0.061
	(0.113)	(0.127)	(0.108)	(0.135)	(0.09)	(0.106)	(0.113)	(0.107)	(0.191)	(0.095)	(0.096)	(0.109)	(0.092)	(0.128)	(0.084)
Df	30.193	5.487	6.464	8.666	7.344	22.723	5.456	6.733	8.601	6.861	15.559	5.247	6.318	10.986	7.964
	(63.56)	(1.997)†	(2.434)†	(5.294)*	(3.702)†	(31.152)	(1.939)†	(2.537)†	(5.680)	(3.293)†	(13.65)	(1.192)*	(1.627)†	(5.725)*	(3.199)†
Diagnostics tests															
LB (8)	17.918*	14.889	11.321	59.333*	6.137	18.328*	14.534	11.234	60.976†	6.425	16.78*	19.12	11.14	60.883†	6.087
LB^2(8)	2.009	4.203	3.5	23.787†	5.507	1.467	4.615	3.847	22.561†	7.112	3.087	4.1	11.348	13.037	5.020
Norm	1.943	168.7	77.263	125.64	15.935	1.408	134.75	55.354	110.95	19.464	6.3	159.33	40.57	90.678	27.692
	(0.379)	(0.00)†	(0.000)†	(0.000)†	(0.000)†	(0.495)	(0.000)†	(0.000)†	(0.000)†	(0.000)†	(0.043)	(0.000)†	(0.000)†	(0.000)†	(0.000)†
Log	883.76	357.94	761.69	753.16	412.62	881.75	357.55	761.57	753.44	412.33	879.08	351.25	754.5	750.37	409.58
AIC	-8.881	-4.664	-6.492	-6.396	-3.473	-8.881	[-4.671]	[-6.507]	[-6.407]	-3.479	[-8.884]	-4.646	-6.473	-6.407	[-3.481]
HQ	-8.820	-4.610	-6.438	-6.348	-3.425	-8.834	[-4.629]	[6.466]	-6.365	-3.438	[-8.857]	-4.622	-6.449	[-6.383]	[-3.458]

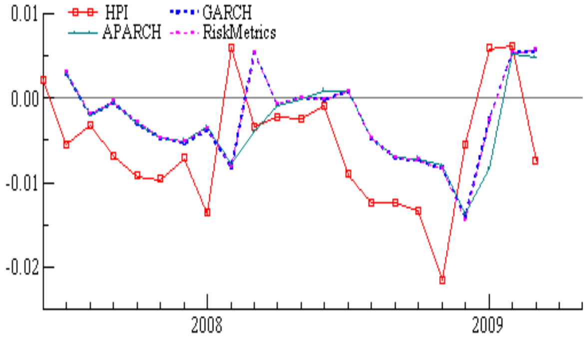
Notes: Figures in parenthesis represent the standard errors of the coefficients in univariate APARCH models. † and * represent statistical significance at 5% and 10%, respectively. *LB* (8) and LB^2 (8) represent the Ljung-Box test statistics at lag 8 for standardized and squared standardized residuals, respectively. Norm and Log stand for Normality and Logliklihood, respectively. AIC and HQ are Akaike Information Criterion and Hannan-Quinn Criterion, respectively. [ ] shows the best fit model.

In the case of CL, due to the misspecification of the conditional mean in the data series, we have to consider higher lagged values for the dependent variables in the mean equation. The Akaike Information Criterion (AIC) statistics suggest that up to two lags should be added in the mean equations and the value of $LB(8)$ and $LB^2(8)$ would then decline, respectively, to 15.21 and 9.08 in the AR (2) - APARCH specification. Similar results are also found in the AR (2) - GARCH and AR (2) - RiskMetrics models. The problem of serial correlation and misspecification in conditional variance equation are therefore eliminated from CL.

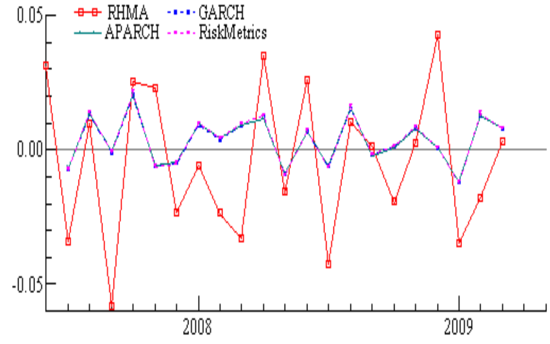
Because the same data sets are used, the results with regard to the significance of the coefficients are almost the same among the three models. However, the one difference among the three models is the specification and the value of coefficients. The Akaike Information Criterion (AIC) and Hannan-Quinn Criterion (HQ) are used to evaluate the in-sample goodness of fit of the models. In accordance with the criterion, the RiskMetrics model provides the best performance in HPI and IL while the conventional GARCH model provides better fit in RHMA and LTP. For the CL, the AIC and HQ criteria do not obtain a consistent result in measuring the goodness of fit among the models. However, it is obvious that the APARCH model provides the poorest performance for the in-sample estimation.

Out-of-sample forecast

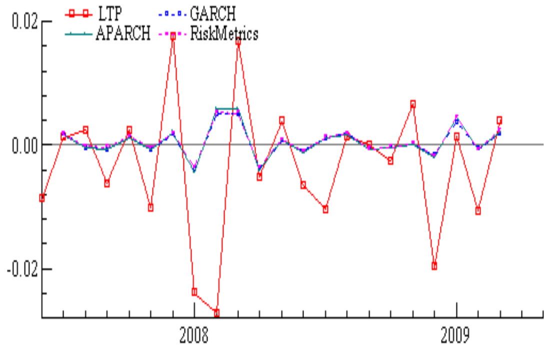
The out-of-sample forecast starts with the estimation of each model using in-sample period data to predict the one-step-ahead (month) volatility forecasts, and the in-sample estimation period is shifted forward by one period for estimation and prediction. This process is repeated N times until the last observation of the forecasting period. The date of the collapse of the Lehman Brothers in mid-September is chosen to distinguish the out-of-sample forecast into the pre-crisis and post-crisis periods. The pre-crisis period for the forecasting analysis is ranged from June 2007 to September 2008, while the post-crisis period is ranged from September 2008 to March 2009. The predicted one-month-ahead volatility is then compared with the realized volatility and all the estimated results are recorded for the models comparison using the statistical tests.



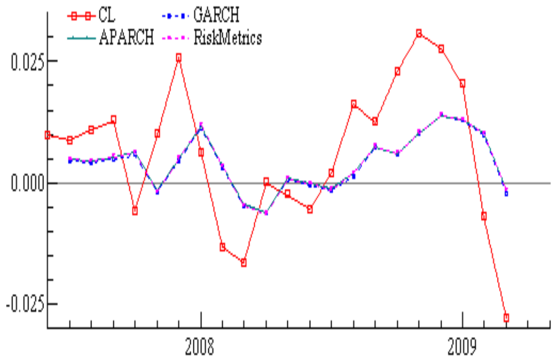
(a) Housing price index



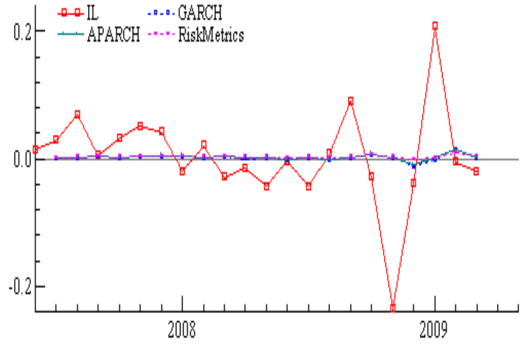
(b) Real home mortgage amount



(c) Loan to price ratio



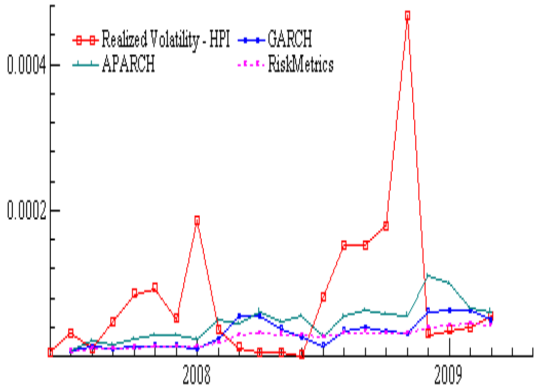
(d) Commercial bank assets – consumer loan



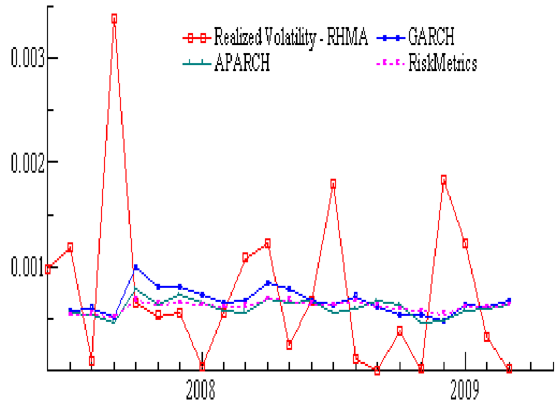
(e) Commercial bank assets- inter-bank loan

Figure 1 Conditional mean forecasts

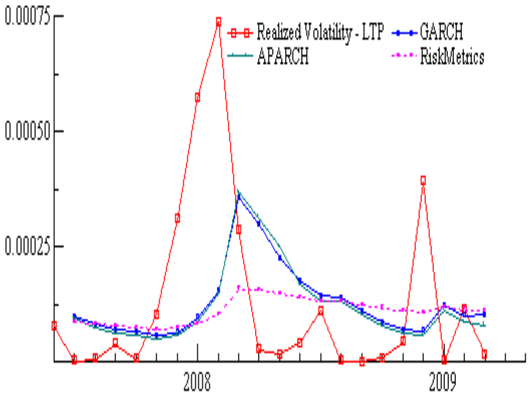
The five portions of Figure 1 illustrate the conditional mean forecast of HPI, RHMA, LTP, CL and IL. Due probably to the similar specification in the mean equation of the three models, the general impression is that their conditional mean forecasts are almost the same. However, there is clearly a significant deviation between the conditional mean forecasts and observed series in all cases, especially after the period of mid-September 2008. Note that the deviation is more serious in the case of CL and IL, as both have declined sharply resulting from the housing market downturn and the delinquent or foreclosure of home mortgage.



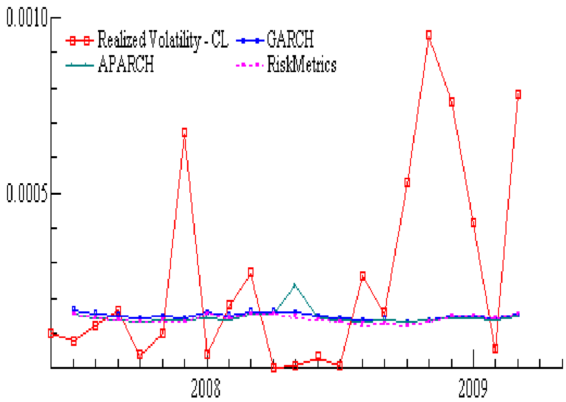
(a) Housing price index



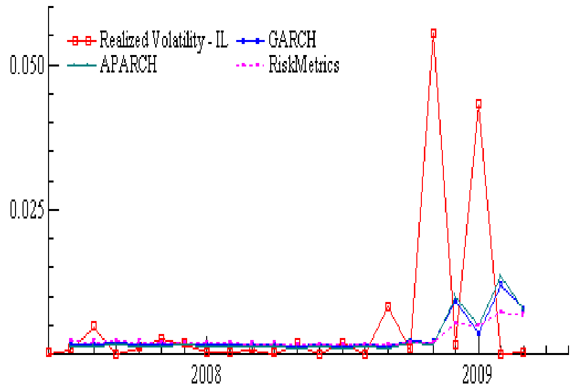
(b) Real home mortgage amount



(c) Loan to price ratio



(d) Commercial bank assets – consumer loan



(e) Commercial bank assets- inter-bank loan

Figure 2 Conditional volatility forecasts

The five portions in Figure 2 present the conditional volatility forecasts compared with the realized volatility for the various models. It is apparent that the prediction from the various models have under-predicted the volatility in each series. In addition, a high deviation between the conditional volatility forecasts and realized volatility could be seen from the beginning of May 2008 in most cases. However, it seems that all models have failed to capture this *ex-post* information available at time *t* to generate a more accurate volatility prediction.

Table 3 shows the actual and relative values of the pre-crisis and post-crisis forecast error statistics on conditional mean for each model across the five error measures. First, consider the results of MAE, RMSE and MAPE statistics. For the HPI, all three statistics indicate that the APARCH model provides the most accurate forecast in both pre-crisis and post-crisis periods. This can be seen from the actual values of APARCH, which is smallest among the three models (0.0051, 0.0067, 0.0062 in the pre-crisis period, and 0.0063, 0.0075, 0.0071 in the post-crisis period). Both the MAE and MAPE statistics suggested that the APARCH model gave, respectively, 15% (1 - 0.8472) and 18% (1 – 0.8196) more accurate forecast than the RiskMetrics model in the pre-crisis period.

Other than HPI, the RiskMetrics model ranked second in accuracy in the case of RHMA in both forecast periods. The APARCH model ranked third, and gave a relative poor performance in both pre-crisis and post-crisis periods among three models even though their relative values are close to each other. For both CL and IL, the RiskMetrics model provided the poorest forecasts among the three models. However, both the difference in the accurate and relative values is extremely small, suggesting that the performance among the three models do not have a significant difference in forecasting the conditional mean of CL and IL.

Table 3 Error statistics from forecasting monthly conditional mean

	(June 2007 – September 2008)						(June 2007 – March 2009)					
	MAE		RMSE		MAPE		MAE		RMSE		MAPE	
	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative
HPI												
APARCH	0.0051	0.8472	0.0067	0.9423	0.0062	0.8196	0.0063	0.9390	0.0075	0.9580	0.0071	0.8568
GARCH	0.0060	0.9884	0.0071	0.9930	0.0074	0.9854	0.0066	0.9881	0.0078	0.9924	0.0082	0.9807
RiskMetrics	0.0060	1.0000	0.0071	1.0000	0.0075	1.0000	0.0067	1.0000	0.0079	1.0000	0.0083	1.0000
RHMA												
APARCH	0.2190	1.0000	0.0271	1.0000	0.0230	1.0000	0.0217	1.0000	0.0265	1.0000	0.0237	1.0000
GARCH	0.2123	0.9694	0.2619	9.6642	0.2191	9.5178	0.0212	0.9770	0.0258	0.9744	0.0228	0.9628
RiskMetrics	0.0212	0.0970	0.0262	0.9675	0.0222	0.9631	0.0213	0.9797	0.0259	0.9762	0.0232	0.9814
LTP												
APARCH	0.0156	0.9911	0.1231	0.9959	0.0117	0.9560	0.0080	1.0000	0.0113	1.0000	0.0081	1.0000
GARCH	0.0157	1.0000	0.1236	1.0000	0.0123	1.0000	0.0078	0.9713	0.0112	0.9911	0.0076	0.9420
RiskMetrics	0.0153	0.9727	0.1230	0.9952	0.0122	0.9927	0.0079	0.9813	0.0112	0.9964	0.0077	0.9445
CL												
APARCH	0.0090	0.9956	0.0104	0.9971	0.0593	0.8736	0.0113	1.0000	0.0131	1.0000	0.0472	0.8868
GARCH	0.0090	0.9934	0.0104	0.9952	0.0596	0.8783	0l0113	0.9938	0.0130	0.9923	0.0473	0.8896
RiskMetrics	0.0090	1.0000	0.0105	1.0000	0.0679	1.0000	0.0113	0.9965	0.0130	0.9931	0.0532	1.0000
IL												
APARCH	0.0328	0.9939	0.0399	0.9945	0.0326	0.9930	0.0495	0.9984	0.0774	0.9990	0.0519	1.0000
GARCH	0.0329	0.9976	0.0400	0.9980	0.0327	0.9976	0.0495	0.9974	0.0773	0.9970	0.0518	0.9977
RiskMetrics	0.0330	1.0000	0.0401	1.0000	0.0328	1.0000	0.0496	1.0000	0.0775	1.0000	0.0512	0.9856

Note: The relative error statistics is measured by expressing the actual statistics as a ratio to the worst performing model for a given error measure.

Table 4 Error statistics from forecasting monthly conditional volatility

	(June 2007 – September 2008)						(June 2007 – March 2009)					
	MAE		RMSE		MAPE		MAE		RMSE		MAPE	
	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative
HPI												
APARCH	0.0001	0.8544	0.0001	0.9223	0.0003	1.0000	0.0001	0.9703	0.0001	0.9206	0.0002	1.0000
GARCH	0.0001	0.9491	0.0001	0.9883	0.0002	0.5926	0.0001	0.9865	0.0001	0.9944	0.0001	0.6087
RiskMetrics	0.0001	1.0000	0.0001	1.0000	0.0001	0.4815	0.0001	1.0000	0.0001	1.0000	0.0001	0.5217
RHMA												
APARCH	0.0006	0.9365	0.0009	0.9923	0.0156	0.9898	0.0006	0.9516	0.0009	0.9884	0.0149	0.9520
GARCH	0.0006	1.0000	0.0009	1.0000	0.0156	0.9900	0.0006	1.0000	0.0009	1.0000	0.0154	0.9827
RiskMetrics	0.0006	0.9206	0.0009	0.9890	0.0157	1.0000	0.0006	0.9355	0.0008	0.9767	0.0156	1.0000
LTP												
APARCH	0.0002	1.0000	0.0002	1.0000	0.0027	1.0000	0.0002	1.0000	0.0002	1.0000	0.0023	1.0000
GARCH	0.0002	0.9793	0.0002	0.9667	0.0025	0.9296	0.0002	0.9974	0.0002	0.9545	0.0023	0.9827
RiskMetrics	0.0002	0.9217	0.0002	0.9613	0.0023	0.8556	0.0002	0.9902	0.0002	0.9545	0.0022	0.9307
CL												
APARCH	0.0001	0.9086	0.0002	1.0000	0.4465	0.7332	0.0002	1.0000	0.0003	1.0000	0.3190	0.7333
GARCH	0.0001	1.0000	0.0002	0.9817	0.6089	1.0000	0.0002	0.9877	0.0003	0.9755	0.4351	1.0000
RiskMetrics	0.0001	0.9872	0.0002	0.9794	0.5405	0.8876	0.0002	0.9842	0.0003	0.9706	0.3862	0.8876
IL												
APARCH	0.0015	0.9321	0.0022	1.0000	0.0189	0.6514	0.0067	1.0000	0.0150	1.0000	0.2552	0.9508
GARCH	0.0015	0.9383	0.0022	0.9776	0.0231	0.7960	0.0067	0.9970	0.0145	0.9726	0.2684	1.0000
RiskMetrics	0.0016	1.0000	0.0022	0.9955	0.0290	1.0000	0.0064	0.9552	0.0147	0.9806	0.1150	0.4286

Note: Same as Table 3.

Table 5 Error statistics from forecasting monthly conditional mean and volatility

	(June 2007 – September 2008)								(June 2007 – March 2009)							
	Conditional mean				Conditional volatility				Conditional mean				Conditional volatility			
	MME(U)		MME(O)		MME(U)		MME(O)		MME(U)		MME(O)		MME(U)		MME(O)	
	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative	Actual	Relative
HPI																
APARCH	0.0067	1.0000	0.0787	1.0000	0.0013	0.5899	0.0043	1.0000	0.0158	0.8502	0.0605	0.9738	0.0035	0.7936	0.0031	1.0000
GARCH	0.0065	0.9701	0.0763	0.9688	0.0019	0.8848	0.0033	0.7685	0.0184	0.9860	0.0617	0.9937	0.0042	0.9610	0.0024	0.7717
RiskMetrics	0.0066	0.9776	0.0767	0.9737	0.0022	1.0000	0.0023	0.5370	0.0186	1.0000	0.0621	1.0000	0.0044	1.0000	0.0017	0.5434
RHMA																
APARCH	0.0231	1.0000	0.1344	0.9754	0.0004	0.6087	0.0168	0.9016	0.0534	1.0000	0.1012	0.9792	0.0073	0.9747	0.0122	0.9041
GARCH	0.0216	0.9321	0.1364	0.9903	0.0005	0.7536	0.0186	1.0000	0.0510	0.9556	0.1024	0.9916	0.0072	0.9521	0.0135	1.0000
RiskMetrics	0.0218	0.9429	0.1378	1.0000	0.0007	1.0000	0.0165	0.8844	0.0507	0.9489	0.1033	1.0000	0.0075	1.0000	0.0119	0.8870
LTP																
APARCH	0.0128	1.0000	0.0774	0.9623	0.0001	0.9091	0.0097	0.9868	0.0273	1.0000	0.0565	0.9660	0.0027	0.9676	0.0072	0.9808
GARCH	0.0105	0.8164	0.0795	0.9891	0.0001	1.0000	0.0098	1.0000	0.0241	0.8820	0.0579	0.9896	0.0026	0.9245	0.0073	1.0000
RiskMetrics	0.0104	0.8133	0.0804	1.0000	0.0001	0.8182	0.0094	0.9522	0.0239	0.8757	0.0585	1.0000	0.0028	1.0000	0.0067	0.9207
CL																
APARCH	0.0352	0.9838	0.0617	0.9981	0.0020	1.0000	0.0069	0.9001	0.0565	0.9902	0.0476	0.9985	0.0066	1.0000	0.0050	0.9024
GARCH	0.0351	0.9796	0.0618	1.0000	0.0016	0.8061	0.0076	1.0000	0.0562	0.9863	0.0477	1.0000	0.0062	0.9376	0.0055	1.0000
RiskMetrics	0.0358	1.0000	0.0609	0.9845	0.0017	0.8878	0.0072	0.9514	0.0570	1.0000	0.0470	0.9866	0.0063	0.9635	0.0053	0.9530
IL																
APARCH	0.0494	0.9932	0.1493	1.0000	0.0027	0.8374	0.0464	0.9036	0.0901	0.9914	0.1176	1.0000	0.0130	1.0000	0.0338	0.8998
GARCH	0.0495	0.9966	0.1484	0.9939	0.0033	1.0000	0.0514	1.0000	0.0906	0.9969	0.1171	0.9957	0.0122	0.9368	0.0375	0.9989
RiskMetrics	0.0497	1.0000	0.1484	0.9940	0.0029	0.8926	0.0507	0.9868	0.0909	1.0000	0.1172	0.9963	0.0108	0.8335	0.0375	1.0000

Note: The relative error statistics is measured by expressing the actual statistics as a ratio to the worst performing model for a given error measure.

Table 4 shows the pre-crisis and post-crisis forecast errors statistics on conditional volatility. The MAE and RMSE statistics suggest that the RiskMetrics model has provided the most inaccurate forecast in both pre-crisis and post-crisis periods. Similar to the conditional mean forecast result shown in Table 3, the MAE statistics suggests that the forecast by the APARCH model is 15% more accurate than the RiskMetrics model. Interestingly, the MAPE statistics indicates that the APARCH is the worst model in predicting the conditional volatility among the three models. In addition, the APARCH model is 51% and 48% less accurate than the RiskMetrics model in pre-crisis period and post-crisis period, respectively.

In measuring the conditional volatility forecast of RHMA, the MAE and RMSE statistics suggest that the RiskMetrics model has performed better than both APARCH and GARCH models. However, the MAPE statistics presents an opposite outcome that the RiskMetrics model provides a less accurate forecast in both pre-crisis and post-crisis periods. In the case of LTP, it is clear that the actual value of the pre-crisis and post-crisis forecast error statistics on conditional volatility is similar in various models. There is no clear distinction between models forecast though all statistics show that the APARCH model has the poorest performance on LTP conditional volatility.

The forecast error statistics shown in Table 3 and Table 4 are based on the assumption of symmetric loss function. It is, however, common in practice that under-prediction and over-prediction are not equally weighted by investors. Table 5 presents the MME statistics that show the over-prediction and under-prediction in the conditional mean and volatility for the pre-crisis and post-crisis periods, respectively. It could significantly be seen that both conditional mean and conditional volatility have generally been over-predicted by all models in all cases as the actual values of MME(O) in both conditional mean and volatility are considerably higher than that of MME(U). Conversely, under-prediction in conditional mean and volatility is common during the post-crisis period.

Due to the downturn of the U.S. economy after the collapse of Lehman Brothers in mid-September 2008, the magnitude of financial market volatility has increased remarkably. One can see from the result of forecast error statistics in Table 3 and Table 4 that the actual value of forecast error in each model is much larger in the post-crisis than in the pre-crisis period. In contrast to the actual value of MME statistics in pre-crisis period shown in Table 5, it is apparently that the actual values of MME(U) of all underlying variables have increased sharply in post-crisis period. The MME(U) statistics in Table 5 that weights under-prediction errors more heavily has shown that the RiskMetrics model has provided the worst conditional mean forecast for HPI, LTP and IL. For the conditional volatility forecast, the RiskMetrics model has also performed poorly in HPI, RHMA and LTP. It implies that a model with a good performance in in-sample analysis may not provide an accurate out-of-sample forecast.

V Conclusion

This paper has provided the in-sample estimation of the conventional GARCH, APARCH and the benchmark Riskmetrics model on real estates related finance series, and evaluated the out-of-sample conditional mean and volatility forecast performance of various models. The date on the collapse of the Lehman Brothers in mid-September is chosen to distinguish the out-of-sample forecast into pre-crisis and post-crisis periods. The empirical results show that the Riskmetrics model has performed satisfactorily in the in-sample estimation but poorly in the out-of-sample forecast. For the post-crisis out-of-sample forecasts, all models have performed poorly in conditional mean and volatility forecast. This result probably is expected. Nonetheless, the 2008 financial crisis has provided a good insight that banks could have under-estimated the potential risk in the sub-prime mortgages. To a large extent, asset returns generally have a random walk feature. An over-reliance on forecasting the future movement of the asset returns is not an appropriate move in reality.

Although standard models have been used in the volatility forecast exercise, the performance of the data prior to the 2008 crisis can provide lessons on the riskiness of real estate finance. Nonetheless, the volatility forecasting analysis using the U.S. real estates related finance data poses challenges on existing methodologies and the use of other suitable proxy variables in forecasting the risk in the real estate market. This paper has probably brought out the problem and left the solution to future studies.

References

Andersen, T. and T. Bollerslev, (1998), "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39, 885-905.

Black, F., (1976), "Studies of stock market volatility changes," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Bollerslev, T., (1986), "Generalized autoregressive conditional heteroskedasiticity," *Journal of Econometrics*, 31(3), 307–327.

Brailsford, T. J. and R. W. Faff, (1996), "An evaluation of volatility forecasting techniques," *Journal of Banking and Finance*, 20, 419-438.

Brooks, C. and G. Persands, (2002), "Model choice and value-at-risk performance," *Financial Analysts Journal*, 58, 87-97.

Christie, A. A., (1982), "The stochastic behavior of common stock variances," *Journal of Financial Economics*, 10, 407–432.

Ding, Z. C., W. J. Granger, and R. F. Engle, (1993), "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83–106.

Engle, R. F., (1982), "Autoregressive conditional heterosedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987–1007.

Engle, R. F., and Bollerslev, T., (1986), "Modeling the persistence of conditional variances," *Econometric Reviews*, 5, 1-50.

Engle, R. F., and V. K. Ng, (1993), "Measuring and testing the impact of news on volatility," *Journal of Finance*, 48, 1749-1778.

Financial Services Authority, (2009), *The Turner Review: A Regulatory Response to the Global Banking Crisis*, London, March.

French, K., G. Schwert, and R. Stambaugh, (1987), "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3–29.

French, Kenneth R., Martin N. Baily, John Y. Campbell, John H. Cochrane, Douglas W. Diamond, Darrell Duffie, Anil K. Kashyap, Frederic S. Mishkin, Raghuram G. Rajan, David S. Scharfstein, Robert J. Shiller, Hyun Song Shin, Matthew J. Slaughter, Jeremy C. Stein, René M. Stulz, (2010), *The Squam Lake Report: Fixing the Financial System*, Princeton: Princeton University Press.

Glosten, L. R., R. Jagannathan, and D. E. Runkle, (1993), "On the relation between expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779–

1801.

J. P. Morgan, (1996), *Riskmetrics Technical Document*, 4th ed., J. P. Morgan, New York.

Markowitz, H., (1952), "Portfolio selection," *Journal of Finance*, 7(1), 77-91.

Nelson, D. B., (1991), "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 59, 347-370.

Sadorsky, P., (2006), "Modeling and forecasting petroleum futures volatility," *Energy Economics*, 28, 467-488.

Schwartz, A. J., (2009), "Origins of the financial market crisis of 2008", *Cato Journal*, 29 (1), 19-23.

Schwert, W., (1990), "Stock volatility and the crash of '87'," *Review of Financial Studies*, 3, 77–102.

Taylor, John B., 2009, *Getting Off Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis*, Stanford: Hoover Institution Press.

Taylor, John B., and John C. Williams, (2009), *A Black Swan in the Money Market*, National Bureau of Economic Research, Working Paper 13943, Cambridge, Mass., April.

Taylor S. J., (1986), *Modelling Financial Time Series*, J. Wiley & Sons, New York.

West, K. D., (1996), "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067-1084.

West, K. D., (2001), "Tests for forecast encompassing when forecasts depend on estimated regression parameters," *Journal of Business and Economic Statistics*, 19, 29-33.

West, K. D., (2006), "Forecast evaluation," in *Handbook of Economic Forecasting*, G. Elliott, C. W. J. Granger, and A. Timmermann, (eds.), North Holland.

Wheaton, W., R. Torto, J. Southard and P. Sivitanides, (1999), "Evaluating risk in real estate," *Real Estate Finance*, 16 (2), 15-22.

Wong, D. K. T., and Kui-Wai Li, (2010), "Comparing the performance of relative stock return differential and real exchange rate in two financial crises," *Applied Financial Economics*, 20 (1/2) January, 137-150.