Estimating Real Estate Value-at-Risk Using Wavelet Denoising and Time Series Model

Kaijian He^{1,2}, Chi Xie¹, and Kin Keung Lai^{2,1}

¹ College of Business Administration, Hunan University, Changsha, Hunan, 410082, P.R. China

paulhekj@gmail.com, xiechi@hnu.cn

² Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
mskklai@cityu.edu.hk

Abstract. As the real estate market develops rapidly and is increasingly securitized, it has become an important investment asset in the portfolio design. Thus the measurement of its market risk exposure has attracted attentions from academics and industries due to its peculiar behavior and unique characteristics such as heteroscedasticity and multi scale heterogeneity in its risk and noise evolution etc. This paper proposes the wavelet denoising ARMA-GARCH approach for measuring the market risk level in the real estate sector. The multi scale heterogeneous noise level is determined in the level dependent manner in wavelet analysis. The autocorrelation and heteroscedasticity characteristics for both data and noises are modeled in the ARMA-GARCH framework. Experiment results in Chinese real estate market suggest that the proposed methodology achieves the superior performance by improving the reliability of VaR estimated upon those from traditional ARMA-GARCH approach.

Keywords: Value at Risk, Real Estate Market, Wavelet Analysis, ARMA-GARCH Model.

1 Introduction

Traditionally the direct investment in the real estate market has been seen as the low risk/low return means of investment and are only allocated the small proportion of the portfolio value [1]. In recent years, the increasing securitization of the real estate industry has offered an important alternative to the direct investment in property. These indirect investment means offer clear advantages over the direct investment counterpart such as significantly higher liquidity level, lower unit costs, lower management and transaction costs, good levels of within property portfolio diversification, and transparent and real time price. Empirical researches also confirm that prices of these securitized real estate vehicles are highly correlated with the value of the underlying properties, thus serves as the alternative to the traditional direct property investment [2].

However, the securitization process of the property investment also introduces additional risks in the investment process. Empirical researches suggest that their prices are not only exposed to the performance of the underlying property values, but also exposed to the volatility in the stock market. Thus they are more volatile than the underlying property market. There is higher level of risks accompanied by higher level of returns with the advent of these new features in the property investment [2].

Therefore, when new features of securitized real estate investment attract large allocation of funds into the real estate investment in the portfolio design, it also demands the appropriate risk measurement and management during the investment process. The more volatile market environment requires more precise measurement and management of the market risk level. However, traditional risk measurement techniques no longer suffice as the property shares increasingly exhibit unique characteristics. On one hand, its movement reflects the underlying property market behavior and is typically auto correlated and mean reverting. On the other hand, it is also influenced by speculative forces in the stock market and incorporates multi scale nonlinear dynamics structure in its risk evolution [3]. Therefore, new risk measurement techniques that are capable of analyzing risk in the multi scale domain is desired to deal with the particular characteristics in the market.

Despite its significance, the literature on risk measurement for real estate investment is surprisingly scarce. This paper attempts to measure the market risk in the real estate market using Value at Risk methodology. Although numerous risk measures emerge over years, none has received the same level of recognition as VaR at both industrial and academic level. As a single summarizing statistics measuring the maximum possible losses over certain investment horizon at the given confidence level, VaR is simple and concise to understand and implement. To estimate real estate VaR at higher accuracy and reliability demanded by investors, this paper proposes the wavelet denoising ARMA-GARCH approach to analyze the multi scale structure of risk evolutions in the real estate market and further estimates VaR at the finer time scale domain.

The contribution of the present paper is two folds: firstly wavelet denoising ARMA-GARCH algorithm is proposed for VaR estimates (WDNVaR). This approach is unique as the wavelet denoising algorithm is treated here as a blind source separation tool that extracts two time series with distinct features from the original time series, i.e. data and noise, for further modeling. Besides, noises are taken into account when measuring the market risk level since it contributes significantly to the overall variation. Secondly the proposed approach is applied to Chinese real estate market to measure significant market risk there, which is largely overlooked issue in the current literature, to the best of our knowledge.

The rest of this paper is organized as follows: the second section reviews the relevant theories. The third section proposes the wavelet denoising ARMA-GARCH VaR estimation methodology. The fourth section conducts empirical studies and analyze results. The fifth section concludes.

2 Relevant Theories

2.1 Value at Risk and Backtesting Theories

VaR measures the maximum level of downside risk of the portfolio at certain confidence level with the given investment time horizon. It computes the maximum loss level as in (1)

$$p\{r_t \leq -r_{VaR}\} = 1 - cl \quad (1)$$

Where r_t denotes the return of the portfolio over the time period t . cl denotes the associated investment confidence level.

There are mainly three approaches to calculate VaR based on historical returns. These include parametric, nonparametric and semi parametric approach. The parametric approach derives the analytical form of risk measures based on the assumed density function. The non parametric approach infers patterns and the risk exposure level from the historical data using empirical data driven approaches. As parametric and non parametric approaches both have their advantages and disadvantages, the semi parametric approaches attract significant attentions recently by complimenting one using the other [4].

Meanwhile, the VaR model adequacy is checked using formal statistical procedures such as Backtesting. There are different types of backtesting procedures developed over years, which mainly concentrate on the following three aspects: the frequency of exceedances, the temporal behavior of exceedances and the size of losses [4]. Among them, Kupiec backtesting procedure is the most widely accepted one and is adopted in this paper to evaluate the model performance. The backtesting process is treated as a series of Bernoulli trials. The frequency of exceedances in these trials should converge to the binomial distribution at infinity. Thus the likelihood ratio test statistics takes the mathematical form as in (2).

$$LR_{uc} = -2\log[(1 - \rho)^{T-N} \rho^N] + 2\log([1 - (\frac{T}{N})]^{T-N} (\frac{T}{N})^N) \quad (2)$$

Where LR_{UC} denotes the test statistics that has an asymptotic $\chi^2(1)$ distribution. N is the number of VaR exceedances. ρ is the probability of an VaR exceedances occurrence. T is the number of trials carried out.

2.2 Wavelet Denoising Theory

Wavelet analysis is a new tool designed for multi scale analysis of signals in the time scale domain. Compared to the frequency only signal projection in the traditional Fourier analysis, wavelet analysis projects signals in the frequency-time domain, thus bringing the capability for more subtle and detailed data analysis. Mathematically wavelet analysis is an orthogonal transform that utilizes a special family of functions called wavelets to approximate and extract the features of interests from the original data [5]. Wavelets are basis functions that satisfy the admissibility and unit energy condition as in (3).

$$C_\varphi = \int_0^\infty \frac{|\varphi(f)|}{f} df < \infty, \int_{-\infty}^\infty |\psi(t)|^2 dt = 1 \quad (3)$$

Where $\varphi(f)$ is the Fourier transform of wavelet $\psi(t)$ in the frequency domain.

The original data series can be orthogonally transformed using wavelet basis functions. It is a decomposition process that projects the original data into the sub data series in the wavelet domain as in (4).

$$W(u, s) = \int_{-\infty}^\infty x(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) dt \quad (4)$$

Where u is the wavelet parameter translating the original wavelet function, and s is the scale parameter dilating the original wavelet function.

The wavelet synthesis process reconstructs the data series using wavelet coefficients in the time scale domain as in (5).

$$x(t) = \frac{1}{C_\psi} \int_0^i nfty \int_{-\infty}^\infty W(u, s) \psi_{u,s}(t) du \frac{ds}{s^2} \quad (5)$$

As data are often contaminated by irrelevant noises, researches on recovering data of interests from noisy data have attracted significant attentions. Traditional approach assumes linear nature of the data during the denoising process and only achieves limited success. Recently nonlinear denoising algorithm that recognizes the nonlinear nature of data and noise has been explored further to improve the quality of data recovered from noisy data. Among numerous denoising algorithm proposed, wavelet analysis is one promising approach [6]. It starts by firstly decomposing the original data series in the time scale domain. The threshold level is selected for each scale following certain threshold selection rules. Then wavelet coefficients smaller than the calculated threshold level are removed while the remaining coefficients are dealt with following certain thresholding strategy. And finally the denoised data series are reconstructed from the denoised sub data series at different scales. In the end, the original data are decomposed into data part and irrelevant noises part.

During the denoising process, both thresholding rules and threshold selection rules are crucial to the appropriate separation of data from noises. There are mainly two thresholding rules in the literature. The hard thresholding rule suppresses only those insignificant data that are less than the chosen thresholds. It retains the rest of the data. Thus it would retain the spike and abrupt changes. The soft thresholding rule subtracts the thresholding value from all data points while suppressing those insignificant data smaller than the chosen thresholds. The soft thresholding rule would leave more smooth and continuous data series for model fitting. Significant efforts have been devoted to develop algorithms for determining the optimal thresholds. These may include numerous threshold selection rules such as universal thresholding, minimax thresholding and Stein's Unbiased Risk Estimate (SURE), etc [6]. The universal thresholding rule sets the upper ceiling boundary for independently and identically distributed noises.

The minimax thresholding rule adopts more conservative approach by choosing the optimal threshold that minimizes the overall mean square error. Thus the abrupt changes from the original data are retained [7]. The SURE chooses the optimal threshold by minimizing the Stein's unbiased estimator of risk and is data adaptive. However, it could be narrow banded around the minimum value range.

3 Wavelet Denoising ARMA-GARCH Approach for Value at Risk Estimates

The traditional ARMA-GARCH approach fits conditional models directly to the original data to obtain parameter estimates for both mean and volatility. It only achieves the moderate level of accuracy since they ignore the underlying risk structure. Besides, different parts of risks may follow different evolution processes, which require the same model with different parameters or different model specifications. The traditional ARMA-GARCH approach is only approximation and offers limited insights into the underlying risk structure and evolution.

To tackle these issues, the wavelet denoising ARMA-GARCH approach is proposed to model the risk evolution at finer scales. Risks are separated into data and noise parts. They are modeled separately with different model specifications and parameters. The rationale behind the proposed wavelet denoising ARMA-GARCH approach is as follows: Since the fundamental and speculation part of real estate risks are driven by completely different factors, they evolve in independent manners and require separate treatments. Thus by separating these two parts and fitting models to them independently using the wavelet denoising algorithm, the fitting accuracy can be improved and more insights can be gained.

The WDNVaR algorithm involves the following steps:

- (1) Wavelet analysis is used to separate the data from irrelevant noises.

$$r_t = r_{data,t} + r_{noise,t} \quad (6)$$

- (2) The mean for both data and noises are modeled by ARMA processes with different parameters.

$$\hat{\mu}_t = a_0 + \sum_{i=1}^m a_i r_{t-i} + \sum_{j=1}^n b_j \varepsilon_{t-j} \quad (7)$$

Where $\hat{\mu}_t$ is the estimate of the conditional mean at time t , r_{t-i} ($i = 1 \dots m$) is the lag m returns with parameter a_i , and ε_{t-j} ($j = 1 \dots n$) is the lag n residuals in the previous period with parameter b_j .

- (3) The conditional mean for the original data series are aggregated from estimates for both data and noises part.

$$\hat{\mu}_{aggregated} = \hat{\mu}_{data} + \hat{\mu}_{noises} \quad (8)$$

- (4) The volatility for both data and noises are modeled by GARCH processes with different parameters.

$$\hat{\sigma}_t^2 = \omega + \sum_{i=1}^p a_i \sigma_{t-i}^2 + \sum_{j=1}^q \beta_j \varepsilon_{t-j}^2 \quad (9)$$

Where $\hat{\sigma}_t^2$ is the estimate of the conditional variance at time t , σ_{t-i}^2 ($i = 1 \dots p$) is the lag p variance with parameter a_i , and ε_{t-j}^2 ($j = 1 \dots q$) is the lag q squared errors with parameter β_j in the previous periods.

(5) And the conditional volatility for the original data series are aggregated from estimates for both data and noises part.

$$\hat{\sigma}_{aggregated}^2 = \hat{\sigma}_{data}^2 + \hat{\sigma}_{nosies}^2 \quad (10)$$

The model order during model fitting process is determined under AIC and BIC minimization principle.

(6) Although the empirical sample distribution could take any arbitrary form, it would converge to the normal distribution with significantly large size of samples and length of time horizon. Thus, when sample size is significantly large and covers significant length of time, VaR can be estimated parametrically as in (11).

$$WDNVaR = -F(a)\hat{\sigma}_{t+1|t} - \hat{\mu} \quad (11)$$

Where $F(a)$ refers to the corresponding quantile (i.e. 95%, 97.5% or 99%) of the normal distribution. $\hat{\sigma}_{t+1|t}$ refers to the estimate of conditional standard deviation at time $t + 1$. $\hat{\mu}$ refers to the estimate of the sample mean.

4 Empirical Studies

The data analyzed in this paper include the closing prices for the Shanghai real estate index. Data are obtained from Global Financial Data, which is one of the major data vendors. The data set includes 3568 observations for Shanghai SE real estate index covering the time period from 30th April, 1993 to 14th December, 2007. Data range covers the period of significant changes and events, such as the residential reform in 1998, the volatile market environment during the transitional period and the late 2007, etc. Thus the proposed model is put to strict test under different market scenarios. 60% of the data set is used as the training set while the remaining 40% is reserved as the test set for model back-testing. The Augmented Dickey Fuller (ADF) test of stationarity shows that the original price series is not stationary. Further autocorrelation and partial autocorrelation function analysis suggest the existence of significant trend factors. Thus the price series are log differenced at the first order to eliminate the trend factors as in $ln \frac{P_t}{P_{t-1}}$.

Some stylized facts can be observed in table 1. There are considerable fluctuations in the Chinese real estate market as suggested by the significant level of standard deviation. The market is also prone to high probability of extreme events occurrences as indicated by significant excess kurtosis. This is further

Table 1. Descriptive Statistics and Statistical Tests

Mean	Standard Deviation	Skewness	Kurtosis	JB Test	BDS test
0.0001	0.0277	1.3826	17.3717	0.001	0

confirmed by the rejection of Jarque-Bera test of normality, which suggests that the phenomenon of fat tail is statistically significant. Meanwhile, the rejection of BDS test of independence suggests that there are nonlinear autocorrelation for the market returns. Thus the risks in the Chinese real estate market are statistically significant. Accurate and reliable risk measures are desired to protect investors’ interests in the risky market environment.

Firstly VaR is estimated using traditional ARMA-GARCH approach. The lag order up to 5 for both ARMA and GARCH processes are tried using the training data. The model order is determined as ARMA(1,1)-GARCH(4,2) under AIC and BIC minimization principle. The rolling window method is used to include new observations made available for each trail. The window length is set to 2139 to cover the relevant information set.

Table 2. ARMA-GARCH VaR for Shanghai SE Real Estate Index

Confidence Levels	Exceedances	Kupiec Test	P Value
99.0%	17	0.4933	0.4825
97.5%	38	0.1489	0.6996
95.0%	64	0.8352	0.3608

As indicated by lower p value, the VaR estimated at 97.5% confidence level is too aggressive. Meanwhile, the estimation at 95% confidence level is too conservative, which leads to too few exceedances. Besides, when it comes to the further performance improvement, this approach leaves little room. Thus the wavelet denoising ARMA-GARCH model is proposed to investigate the underlying risk structure and further improve the accuracy and reliability of VaR estimated.

Four parameters need to be set when estimating VaR based on the wavelet denoising approach, including the decomposition level, the wavelet family, the thresholding rule and the threshold selection rule. The decomposition level is set to 1. The wavelet families tried includes Haar, Daubechies 2 and Coiflets 2. The thresholding rule tried include hard and soft thresholding rule. The threshold selection algorithms attempted include SURE, minimax and universal threshold selection algorithm. Firstly the wavelet denoising algorithm is used to denoise the original data series and analyze the underlying risk structure. The standard deviations for both denoised and noise data are calculated to investigate the impacts of both data and noises on the total risk level. The ARMA-GARCH model is fitted to the denoised data and noise separately to determine the model order. In table 3, the standard deviation is shown as the number outside the

Table 3. ARMA-GARCH Model Order with different Thresholding and Threshold Selection Rules

Wavelets	Hard	Soft	Hard	Soft	Hard	Soft
	Heursure	Heursure	Minimaxi	Minimaxi	Universal	Universal
Data	(2,3,1,1)	(5,4,1,4)	(5,4,1,5)	(3,4,4,4)	(5,5,4,4)	(4,5,4,4)
(Haar)	0.0273	0.0240	0.0255	0.0216	0.0238	0.0205
Noise	(1,2,5,4)	(1,1,4,4)	(1,2,5,4)	(1,1,4,4)	(1,1,4,4)	(1,1,4,4)
(Haar)	0.0046	0.0082	0.0108	0.0140	0.0141	0.0164
Data	(1,3,5,2)	(1,1,1,1)	(5,5,2,2)	(2,4,1,4)	(2,5,4,3)	(5,4,4,4)
(Db2)	0.0271	0.0237	0.0255	0.0216	0.0239	0.0206
Noise	(1,4,4,5)	(4,5,3,2)	(2,2,4,5)	(4,5,2,2)	(4,5,4,4)	(2,4,3,4)
(Db2)	0.0055	0.0092	0.0107	0.0140	0.0140	0.0163
Data	(3,4,5,2)	(4,5,1,2)	(2,5,1,4)	(4,4,1,2)	(3,5,3,1)	(4,1,4,2)
(Coif2)	0.0273	0.0237	0.0253	0.0211	0.0235	0.0201
Noise	(5,3,1,1)	(1,5,4,2)	(4,5,1,1)	(1,5,4,2)	(3,5,1,5)	(1,5,2,1)
(Coif2)	0.0049	0.0087	0.0113	0.0146	0.0146	0.0172

parenthesis. The number pair (r,m,p,q) shown within the parenthesis correspond to the optimized model order in ARMA(r,m)-GARCH(p,q) specification.

Experiment results in table 3 confirm that the standard deviation for noise data series are significant and non negligible during the modeling process. It could be as large as half of the standard deviation of data series in some circumstances. That means that noise data contributes up to 20% of the total variation in the data. Thus previous approaches ignoring the significant impact of noise data lead to significantly downward biased risk estimates. In contrast, risk estimates for noise data constitute important part of total risk estimates.

Since the typical physical real estate market is usually more stable than the securitized market, the fundamental part of the risk is supposed to be smaller than the speculation part of the risk [?]. Thus it can be observed from experiment results that the noise data correspond to the fundamental part of market risks since it contributes less to the total variation in the long term. Meanwhile the denoised part corresponds to the speculation part of the market risk since they contribute most to the total variation. The securitized real estate market is much more volatile than the traditional one. This findings offer potentially useful information and freedom to investors when estimating the market price and risk level. Based on the analysis, the securitized real estate market is at least 50% more volatile than the traditional one.

Experiment results in table 3 also show that the model order varies when WDNVaR is tried using different parameters. Thus to determine the best set of parameters for improving model's forecasting accuracy, the AIC and BIC information criteria are introduced into the modeling process. Experiment results in

table 3 also shows clearly that data and noise follow their own path of evolutions and they are both non negligible during modeling process. This finding supports the modeling methodology proposed in this paper that treats data and noise separately with emphasis on their own characteristics. It also implies that the current denoising algorithms are not solely reliable as they ignore the potential significant impact of the noise part of the data, determine the separation of data and noise in a largely arbitrary manner and lacks the theoretical guidance as to the suitability of the parameters chosen.

Based on the previous findings, the WDNVaR is used to estimate VaR in Chinese real estate market. Through trial and error method, the decomposition level is set to 1 as it shows the best performance among decomposition levels tested up to 6. The wavelet family chosen is Haar. The universal threshold selection rule is used during the denoising process while the noises are processed following hard thresholding method. The model order for fitting denoised data is determined as ARMA(5,5)-GARCH(4,4). The model order for noise data is determined as ARMA(1,1)-GARCH(4,4).

Table 4. WDNVaR(Haar,1) for Shanghai SE Real Estate Index Confidence Level

Confidence	Levels Exceedances	Kupiec	Test P Value
99.0%	17	0.4933	0.4825
97.5%	36	0.0026	0.9595
95.0%	67	0.2912	0.5895

As shown in the experiment results in table 4, the reliability of VaR estimated based on the wavelet denoising approach improves considerably. The P value improves at both 95% and 97.5% confidence level and stays the same at 99% confidence level, during the model backtesting process. The performance improvement is attributed to the separation of data from noises during the model fitting process. Thus, the models fitted to both data and noise series can better capture the essential factors governing the risk evolution and lead to higher goodness of fit.

5 Conclusions

This paper proposes the wavelet denoising ARMA-GARCH approach for estimating VaR. Compared to traditional approaches that recover signals from noisy data, the wavelet denoising ARMA-GARCH model recognizes the contributions of both data and noise to the overall market risk level. The denoised data are linked to the speculation part of the real estate risk while the noise part is linked to the fundamental part of the market risk. Both data and noises are separated by the utilization of wavelet analysis for further model fitting. Recognizing their distinct features, they are further modeled using ARMA-GARCH specification

with different model specifications and parameters. Since empirical studies suggest that Chinese real estate market represents a considerable risky environment, the proposed wavelet denoising ARMA-GARCH approach is applied to Chinese real estate market for market risk measurement. Experiment results confirm the superior performance of the proposed model against traditional ARMA-GARCH approach, i.e. the reliability of VaR estimated improves considerably upon those from the traditional ARMA-GARCH approach.

Acknowledgements

The work described in this paper was supported by a grant from the National Social Science Foundation of China (SSFC No.07AJL005) and a Research Grant of City University of Hong Kong (No. 9610058).

References

1. Glascock, J.: Market conditions, risk, and real estate portfolio returns: Some empirical evidence. *The Journal of Real Estate Finance and Economics* 4(4), 367–373 (1991)
2. Ball, M., Lizieri, C., MacGregor, B.D.: *The economics of commercial property markets*, pp. 367–384, Routledge, London, GB98Z0620 bnb 2516 Michael Ball, Colin Lizieri and Bryan D. MacGregor. 24cm. Bibliography: Includes index (1998)
3. Scott, L.O.: Do prices reflect market fundamentals in real estate markets? *The Journal of Real Estate Finance and Economics* 3(1), 5–23 (1990)
4. Dowd, K.: *Measuring market risk*. Wiley finance series. pp. 341–354, Wiley, Chichester (2002) (GBA364768 bnb 2758 Kevin Dowd. ill. 25 cm. + 1 CD-Rom (4 3/4 in.) Includes bibliographical references and indexes)
5. Percival, D.B., Walden, A.T.: *Wavelet methods for time series analysis*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge (2000) (GBA102097 bnb 2629 Donald B. Percival and Andrew T. Walden. ill. 26cm. Includes bibliographical references and index)
6. Donoho, D.L., Johnstone, I.M.: Adapting to unknown smoothness via wavelet shrinkage. *Journal Of The American Statistical Association* 90(432), 1200–1224 (1995)
7. Donoho, D.L., Johnstone, I.M.: Minimax estimation via wavelet shrinkage. *Annals Of Statistics* 26(3), 879–921 (1998)