Smooth Transition CARR Models

Ray Y. Chou
Institute of Economics Academia Sinica
rchou@econ.sinica.edu.tw

and

Jiaquan Yao
Jinhe Economic Center, X'ian Jiao Tong University
jqyao@163.com

First draft: January 4, 2007
This draft: April 30, 2007

Abstract

Motivated by smooth transition GARCH and smooth transition ACD models, we add smooth transition mechanism into the conditional mean equation of the Conditional AutoRegressive Range (CARR) model of Chou (2005) and propose a new range-based volatility model: the smooth transition CARR model (henceforth STCARR). We derive a misspecification test statistics to test the CARR model against its smooth transition counterpart using the Taylor expansion method suggested by Meitz and Teräsvirta (2006) and investigate how our test statistics behave in finite samples by Monte Carlo simulations. In application part, we find significant improvements of fitness in smooth transition CARR models upon the associated CARR models for the Nikkei 225 index series.

JEL codes: C22, C51, C52
Keywords: volatility, CARR, Smooth Transition CARR

1. Introduction

Time series modelling of financial market volatility has been a central theme in empirical finance. Accurate modeling and forecasting of volatility is important for portfolio analysis, risk management and derivatives pricing. The observation that many financial series exhibit volatility clustering has led to the development of a great many time series methods. The most popular approaches are GARCH family of models.

The ARCH model introduced in Engle (1982) allows the conditional variance change over time as a function of lagged squared residuals leaving the unconditional variance constant. Bollerslev (1986) proposes the Generalized ARCH (GARCH) model by adding lagged conditional variances in the conditional variance equation. There are a large number of extensions of the original GARCH model. For a critical review, see Bollerslev, Chou, and Kroner (1992), and more recently Engle (2002).

An important generalization of the ARCH model is to allow for regime-switching behavior. One of these regime-switching models is the Markov regime switching model, which was proposed by Hamilton and Susmel (1994), Cai (1994), and Dueker (1997). They use Markov regime switching models to deal with the changes in regime in GARCH specifications. In their papers, the conditional variances may structurally change among a discrete number of regimes with the transition between regimes governed by an unobserved Markov chain. These models are good at capturing sharp volatility, which may not be consistent with the real world.

The common characteristic shared by the GARCH-type models is the existence of only two regimes: low and high volatilities, which are triggered by positive and negative shocks, respectively. In this sense, all are threshold models where the threshold is known and is equal to zero.

A decision making agent may change his behavior at a given point over time based on the available information. However, the economy consists of a great number of agents whose behavior may switch sharply but not simultaneously. Hence smooth transition models may be more realistic than Markov regime switching models in providing a more realistic representation of aggregate behaviors in economies. Many works apply smooth transition mechanism from the mean equation to the variance equation in order to describe the smooth changes in parameters against the assumption of parameter constancy in GARCH specification. For example, smooth transition GARCH models are discussed in Hagerud(1997), Gonzalez-Rivera(1998), Anderson, Nam, and Vahid (1999). These models allow parameters in the variance equation to vary over time as continuous functions of a transition variable.

The range, defined as the difference between the highest and lowest log security prices over a fixed sampling interval, is a viable and much more efficient volatility proxy. The literature supporting this fact includes Parkinson (1980), Garman and Klass (1980), Rogers and Satchell (1991), Gallant,Hsu, and Tauche (1999), Yang and Zhang (2000), Alizadeh, Brandt, and Diebold (2002), Chou (2005), Brandt and Jones (2006), among others. In particular, Chou (2005) combines range theory with GARCH model and propose a dynamic volatility model: the Conditional AutoRegressive Range (henceforth CARR) model. It is shown in Chou (2005) that the range can be used as a measure of volatility and the range-based CARR model performs better than the return-based GARCH model in forecasting volatilities of S&P500 stock index. In CARR model, the evolution of the conditional range is specified in a fashion similar to the conditional variance models as in GARCH and is very similar to the ACD model of Engle and Russell(1998) for durations between trades. In fact, they are all special cases of the Multiplicative-Error-Model of Engle (2002).

Recently smooth transition ACD model is introduced by Meitz and Teräsvirta(2006), its main properties are analyzed and it serve as alternative in the tests of linearity. Motivated by smooth transition GARCH and smooth transition ACD models, we add smooth transition mechanism into the conditional mean equation of the CARR model and propose a new range-based volatility model: the smooth transition CARR model (henceforth STCARR).In our model, the conditional mean of range is a non-linear function of lagged ranges and lagged conditional mean of ranges via a transition function. Two commonly used smooth transition functions are the logistic function and exponential function.

The smooth transition CARR model is an extension of the two-regime volatility, since it allows intermediate states of regimes. The moving among regimes is dictated by an observable transition variable that belongs to the history of the process. It also nests a threshold specification, since for certain parameter values, smooth transition CARR models collapse to threshold models.

Testing for the existence of a smooth-transition mechanism in CARR models presents similar problems to those encountered in the smooth transition GARCH and smooth transition ACD models. Under the null hypothesis, there are nuisance parameters that are not identified, they exist only under the alternative. Since these parameters cannot be estimated under the null hypothesis, the standard asymptotic theory does not apply. There are several ways to formulate the null hypothesis of interest. In the context of this paper, we use the Taylor expansion method suggested in Meitz and Teräsvirta (2006) to solve the non-identification problem.

The rest of this paper is organized as follows. We define the STCARR model and discuss its properties in Section 2. We derive a misspecification test in Section 3 and investigate how our test statistics behave in finite samples by Monte Carlo simulations in Section 4. In Section 5 we apply the STCARR model to the daily data of Nikkei 225 index. Finally, we conclude in Section 6.

2. The model

Parkinson (1980) forcefully argues and demonstrates the superiority of using range as a volatility estimator as compared with standard methods. Let P_t be the natural logarithm of price of a speculative asset, and let P_t^{high} and P_t^{low} be the high and low prices. Consequently, in our paper the range can be defined as

$$R_t = 100 * (\ln(P_t^{high}) - \ln(P_t^{low})) \tag{1}$$

Note that R_t is always positive when high and low prices are not equal.

Following Chou (2005), the class of exponential CARR models is defined as follows:

$$R_t = \lambda_t \varepsilon_t$$
$$\lambda_t = w + \sum_{i=1}^{q} \alpha_i R_{t-i} + \sum_{j=1}^{p} \beta_i \lambda_{t-i} \tag{2}$$
$$\varepsilon_t \sim i.i.d. \exp(1)$$

where R_t is the range. λ_t is the conditional mean of the range based on all information up to time t. The disturbance term ε_t is assumed to be distributed with an exponential distribution with a unit mean. The (w, α_i, β_i) coefficients in the conditional mean equation are all positive to ensure positivity of λ_t. The exponential distribution assumption is useful since the estimation of it is simple and also it yields consistent parameter estimates even if the true distribution is not exponential. For discussions of this property of Quasi Maximum Likelihood Estimation, see Engle (2002).

The motivation for smooth transition CARR models comes form the smooth transition GARCH and smooth transition ACD literature mentioned at section 1. Smooth transition CARR(STCARR) model is defined as follows:

$$R_t = \lambda_t \varepsilon_t$$
$$\lambda_t = w + \sum_{i=1}^{q} \alpha_i R_{t-i} + \sum_{j=1}^{p} \beta_j \lambda_{t-j} + \sum_{i=1}^{q} \alpha_i^* R_{t-i} F(\ln R_{t-i}; \gamma, \mathbf{c}) \tag{3}$$
$$\varepsilon_t \sim i.i.d. \exp(1)$$

where,

$$F(\ln R_{t-i}; \gamma, \mathbf{c}) = (1 + \exp\{-\gamma \prod_{k=1}^{K} (\ln R_{t-i} - c_k)\})^{-1}$$
$$\text{with } c_1 \le \ldots \le c_K, \gamma > 0. \tag{4}$$

In practice, K=1 or 2, i.e.,

$$F(\ln R_{t-i};\gamma,\mathbf{c}) = (1+\exp\{-\gamma(\ln R_{t-i} - c_1)\})^{-1} \quad \text{with } \gamma > 0, \tag{5}$$

or $F(\ln R_{t-i};\gamma,\mathbf{c}) = (1+\exp\{-\gamma(\ln R_{t-i} - c_1)(\ln R_{t-i} - c_2)\})^{-1}$ with $\gamma > 0, c_1 \le c_2$.

Eq. (4) reduces to a standard logistic function when K = 1. In this case, if the smoothness parameter $\gamma \to \infty$, Eq. (4) becomes a step function. In specific words, $F(\ln R_{t-i};\gamma,\mathbf{c})$ =1 for $\ln R_{t-i} \ge c_1$ and $F(\ln R_{t-i};\gamma,\mathbf{c})$ =0 for $\ln R_{t-i} < c_1$. When K = 2, the minimum value of $F(\ln R_{t-i};\gamma,\mathbf{c})$ lies between 0 and 1/2 while $F(\ln R_{t-i};\gamma,\mathbf{c}) \to 1$ if $\ln R_{t-i} \to \pm\infty$. In particular, when $\gamma \to \infty$ with K= 2, Eq. (4) turns out to be a double-step function in which $F(\ln R_{t-i};\gamma,\mathbf{c})$ = 0 for $c_1 \le \ln R_{t-i} \le c_2$ and $F(\ln R_{t-i};\gamma,\mathbf{c})$ = 1 otherwise.

The conditional mean of range is a non-linear function of lagged ranges and lagged conditional mean of ranges via a logistic transition function. The logistic function is defined on the whole real axis, whereas the potential transition variable R_{t-i} only takes positive values. So we use $\ln R_{t-i}$ as the transition variable. In order to capture the effects of very high and low volatilities we concentrate on the choice K=2, which allows these extreme volatilities to have an impact different from the one of the more average volatilities. For the smooth transition CARR to be defined it is required that at least one of α_i (or α_i^*) is larger than zero.

Under the assumption for the disturbances follow an exponential distribution with unit mean, the log likelihood function can be written as

$$L = -\sum_{t=1}^{T}[\ln(\lambda_t) + \frac{R_t}{\lambda_t}] \tag{6}$$

The smooth transition CARR model is closely related to the threshold CARR model. We restrict our comparison between the smooth transition CARR(1,1) model and the threshold CARR(1,1) model which is defined as follows.

$$\begin{aligned} R_t &= \lambda_t \varepsilon_t \\ \lambda_t &= w^j + \alpha_1^j R_{t-1} + \beta_1^j \lambda_{t-1}, \quad r_{j-1} \le R_{t-1} < r_j \\ \varepsilon_t &\sim i.i.d. \exp(1) \end{aligned} \tag{7}$$

where $j=1,...,J$,and J is the number of different regimes and $0=r_0<r_1<...<r_J=\infty$ are the threshold values. The parameter values w^j , α_1^j , β_1^j vary according to different regimes. With the restrictions $w^1=w^2$ and $\beta_1^1=\beta_1^2$, the smooth transition CARR(1,1) model with K=1 as $\gamma\to\infty$ can reduce to two-regime threshold CARR(1,1) model. Similarly, with the restrictions $w^1=w^2=w^3$, $\beta_1^1=\beta_1^2=\beta_1^3$ and $\alpha_1^1=\alpha_1^3$ the smooth transition CARR(1,1) model with K=2 as $\gamma\to\infty$ can collapse to three-regime threshold CARR(1,1) model.

3. Misspecification test

We now consider testing the CARR model against its smooth transition CARR model with K=2. It is seen that model (3) is only identified under the alternative. For example, when γ = 0, parameters $\alpha_i^*, i = 1$, ... ,q , as well as c_k , $k = 1$,...,K ,are not identified. There are three approaches to test when some parameters are identified only under the alternative. First one is discussed in Hansen(1996). He studies the (non-standard) asymptotic distribution theory for such tests, and develops a procedure to approximate these distributions by simulation. Second testing procedure is suggested by Davies (1977,1987) and carried out by fixing unidentified parameters. In order to derive an easily applicable misspecification tests, we use the method proposed by Meitz and Teräsvirta (2006). In their approach the identification problem is solved by approximating the transition function with its first-order Taylor expansion around $\gamma = 0$. This will lead to an approximate alternative, which is free of nuisance parameters under the null hypothesis. For the similarity between STCARR and STACD, the procedure we derive the test statistic is very similar to their procedure and we make some necessary modifications.

Using Taylor's expansion we can the following result

$$\begin{aligned} F(\ln R_{t-i};\gamma,c) &\approx F(\ln R_{t-i};0,c) + \frac{\partial F(\ln R_{t-i};0,c)}{\partial \gamma}(\gamma - 0) \\ &\approx 0 + \frac{1}{4}\gamma \prod_{k=1}^{K}(\ln R_{t-i} - c_k) \\ &\approx \frac{1}{4}\gamma \prod_{k=1}^{K}(\ln R_{t-i} - c_k) \\ &\approx \sum_{l=0}^{K} \gamma \widetilde{c}_l (\ln R_{t-i})^l \end{aligned} \tag{8}$$

Applying Eq. (8) to Eq. (3) we can get

$$\begin{aligned} \lambda_t &= w + \sum_{i=1}^{q}\alpha_i R_{t-i} + \sum_{j=1}^{p}\beta_j \lambda_{t-j} + \sum_{i=1}^{q}\alpha_i^* R_{t-i} \sum_{l=0}^{K}\gamma \widetilde{c}_l (\ln R_{t-i})^l \\ &= w + \sum_{i=1}^{q}(\alpha_i + \alpha_i^* \gamma \widetilde{c}_0) R_{t-i} + \sum_{j=1}^{p}\beta_j \lambda_{t-j} + \sum_{i=1}^{q}\sum_{l=1}^{K}\gamma \alpha_i^* \widetilde{c}_l R_{t-i} (\ln R_{t-i})^l \end{aligned} \tag{9}$$

Renaming parameters yields the following conditional mean of the alternative

$$\lambda_t = w + \sum_{i=1}^{q}\alpha_i R_{t-i} + \sum_{j=1}^{p}\beta_j \lambda_{t-j} + \sum_{i=1}^{q}\sum_{l=1}^{K} e_{il} R_{t-i} (\ln R_{t-i})^l \tag{10}$$

Using Eq.(10) we have transformed the original testing problem into the problem of testing the CARR(p,q) against the approximate alternative

$$
\begin{aligned}
&R_t = (\lambda_t + \varphi_t)\varepsilon_t \\
&\lambda_t + \varphi_t = w + \sum_{i=1}^{q}\alpha_i R_{t-i} + \sum_{j=1}^{p}\beta_j(\lambda_{t-j} + \varphi_{t-j}) + \sum_{i=1}^{q}\sum_{l=1}^{K} e_{il} R_{t-i}(\ln R_{t-i})^l \\
&\varphi_t = \sum_{j=1}^{p}\beta_j \varphi_{t-j} + \sum_{i=1}^{q}\sum_{l=1}^{K} e_{il} R_{t-i}(\ln R_{t-i})^l \\
&\varepsilon_t \sim i.i.d.\exp(1)
\end{aligned}
\tag{11}
$$

Model (11) reduces to the null model when e_{il} = 0 for $i = 1,...,q$ and $l = 1,...,K$, and there are no unindentified parameters under the null. So we can use the theorem 2 in appendix A to derive the test statistic, which is given in the following corollary.

Corollary1: Considering the model(11) and denote $\theta_1 = (w, \alpha_1, ..., \alpha_q, \beta_1, ..., \beta_p)'$ and $\theta_2 = (vecE)$,where E= $[e_{il}]$, $i = 1,...,q$, $l = 1,...,K$,are ($q \times K$)matrices, and the vec-operator stacks the columns of the matrix. Moreover, denote $J = vec(X_t)$, where $X_t = [R_{t-i}(\ln R_{t-i})^l]$, $i = 1,...,q$, $l = 1,...,K$,are also ($q \times K$) matrices. Under the null hypothesis $H_0 : \theta_2 = 0$ .let

$$
\begin{aligned}
&\hat{\mathbf{a}}_t = \frac{1}{\hat{\lambda}_t}\frac{\partial \hat{\lambda}_t}{\partial \theta_1} = \hat{\lambda}_t^{-1} * (1, R_{t-1}, ..., R_{t-q}, \hat{\lambda}_{t-1}, ..., \hat{\lambda}_{t-p})' + \hat{\lambda}_t^{-1}\sum_{j=1}^{p}\hat{\beta}_j \frac{\partial \hat{\lambda}_{t-j}}{\partial \theta_1} \\
&\hat{\mathbf{b}}_t = \frac{1}{\hat{\lambda}_t}\frac{\partial \hat{\varphi}_t}{\partial \theta_2} = \hat{\lambda}_t^{-1} * vec(X_t) + \hat{\lambda}_t^{-1}\sum_{j=1}^{p}\hat{\beta}_j \frac{\partial \hat{\varphi}_{t-j}}{\partial \theta_2}
\end{aligned}
\tag{12}
$$

the statistics thereby can be computed by

LM= $T(SSR_0 - SSR_1)/SSR_0$ (13)

which has an asymptotic χ^2 distribution with qK degrees of freedom the F-version test statistic of the statistic (13) is

$$
F = \frac{(SSR_0 - SSR_1)/qK}{SSR_1/(T - p - q - 1 - qK)} \tag{14}
$$

has an asymptotic F distribution with (qK, $T - p - q - 1 - qK$) degrees of freedom.

For the convenience of performing a Monte Carlo experiment to evaluate the finite sample properties of the test statistics. We derive a test statistic to test the CARR(1,1) model against smooth transition CARR(1,1) model with K=2, which is the special case of statistic(13). The proof is given in Appendix B. In the programming of constructing the simulation experiment, we use LM-version test statistic for the reason that there is not distinct different between LM-version statistic and F-version statistic for the outcome in our programming.

4. Monte Carlo simulations

4.1 Size simulations

The above distribution theory is asymptotic, and we have to investigate how our test statistics behave in finite samples by simulations. We begin with size simulations the data generating process we use has the following form.

$$R_t = \lambda_t \varepsilon_t$$
$$\lambda_t = 0.1 + 0.1R_{t-1} + 0.8\lambda_{t-1} \qquad (15)$$
$$\varepsilon_t \sim i.i.d.\exp(1)$$

The length of the generated time series are 500 and 1500 observations after removing the first 500 observations from the beginning of the series to avoid initialization effects. The number of replications is set to 2000. In each replication, a CARR(1,1) model is estimated and then evaluated using the linearity test against smooth transition CARR with K=2.

Results of the experiment are presented in table 1. The empirical size of the test is close to the nominal size for both samples. In the last two columns, we report the empirical mean and the empirical variance of the test that is consistent with the theoretical mean and variance of a χ^2 distribution with two degree of freedom. The theoretical mean and variance of χ^2 distribution with two degree of freedom are 2 and 4 respectively.

Table 1
Size of the LM test

Model under H_0:CARR(1,1)	sample size	Nominal size			χ^2_2	
		10%	5%	1%	mean	variance
$w = 0.1, \alpha_1 = 0.1,$ $\beta_1 = 0.8$	500	9.15	4.4	0.9	1.918	3.807
	1500	10	4.65	0.85	1.966	3.802

4.2 Power simulations

In the power simulation, we consider two alternative data generating processes. The first one is smooth transition CARR(1,1) specification given by

$$
\begin{aligned}
&R_t = \lambda_t \varepsilon_t \\
&\lambda_t = w + \alpha_1 R_{t-1} + \beta_1 \lambda_{t-1} + \alpha_1^* R_{t-1} F(\ln R_{t-1}; \gamma, \mathbf{c}) \\
&F(\ln R_{t-1}; \gamma, \mathbf{c}) = (1 + \exp\{-\gamma(\ln R_{t-1} - c_1)(\ln R_{t-1} - c_2)\})^{-1} - \frac{1}{2} \qquad (16) \\
&\varepsilon_t \sim i.i.d.\exp(1)
\end{aligned}
$$

As the second alternative model we examine a three-regime threshold CARR(1,1) model given by

$$
\begin{aligned}
&R_t = \lambda_t \varepsilon_t \\
&\lambda_t = \begin{cases} 0.05 + 0.20 R_{t-1} + 0.85 \lambda_{t-1} & for\, 0 < R_{t-1} < 0.25 \\ 0.10 + 0.05 R_{t-1} + 0.90 \lambda_{t-1} & for\, 0.25 \le R_{t-1} < 1.5 \\ 0.20 + 0.03 R_{t-1} + 0.80 \lambda_{t-1} & for\, 1.5 \le R_{t-1} < \infty \end{cases} \qquad (17) \\
&\varepsilon_t \sim i.i.d.\exp(1)
\end{aligned}
$$

The motivation of threshold CARR(1,1) model comes from Zhang, Russell, and Tsay (2001). This model is not a special case of our STCARR model. The first regime is an explosive one. The other two regimes are stable, the middle one being more persistent than the third one.

The length of our sample sizes are 500, 1500 and 4500 after discarding the first 500 observations from the beginning of each generated series to eliminate the effects of the initial values. The number of replications is 2000. For each replication, a CARR(1,1) model is fitted to each series. The diagnostic test performed is the test of no smooth transition CARR with K=2. Since our size simulations indicated that the tests in question are well-sized, the power results are not size-adjusted.

In table 2 we report the power of the test when the alternative data generating process is the smooth transition CARR(1,1) given in Eq.(16).

Table 2

Power of the LM test

Model under H_1: smooth transition CARR(1,1)	sample size	Rejection frequencies (%)		
		Nominal size		
		10%	5%	1%
$w=0.1, \alpha=0.1,$ $\beta=0.8, \alpha^*=0.1,$ $c_1=-0.5, c_2=2,$ $\gamma=1$	500	11.6	6.55	1.6
	1500	27.85	18.5	8.75
	4500	67.2	57.5	37.75
	7500	87.85	81.6	65.45
$w=0.1, \alpha=0.1,$ $\beta=0.8, \alpha^*=0.1,$ $c_1=-0.5, c_2=2,$ $\gamma=10$	500	14.25	7.9	2.15
	1500	31.65	22.5	10.45
	4500	67.75	58.6	42
	7500	86.4	81.2	68.6

Table 3 we present the power results for the threshold CARR(1,1) model given in Eq.(17). The test have very good power for 4500 observations, and low power for 500 observations and 1500 observations.

Table 3

Power of the LM test

Model under H_1: Threshold CARR(1,1)	sample size	Rejection frequencies (%)		
		Nominal size		
		10%	5%	1%
The specification given in Eq.(17)	500	29.3	19.9	7.95
	1500	59.4	46.5	24.85
	4500	97.2	94.95	84.85

5. Applications

We estimate smooth transition CARR model to the daily data of Nikkei 225 index. The sample period we use is from 7 October 1991 to 23 August 2005. The data set consists of 3418 observations. These data can be downloaded from the website “finance.yahoo.com”. The daily range is calculated using Eq. (1). Figure 1 presents the plots of the daily range of Nikkei 225 index.

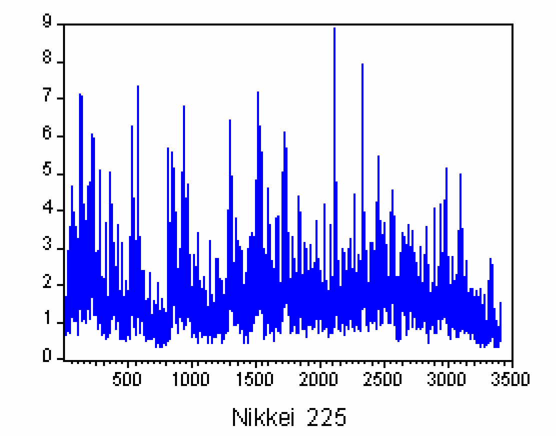

Figure 1: Daily Range Series for Nikkei 225 index over the Period 7 October 1991 to 23 August 2005

The basic statistics of the series is reported in Table 4. The Jarque-Bera statistics for range series is 7580.7 with p-value 0.000 indicating the series departure strongly from normal distribution. But the Jarque-Bera statistics for log-range series is 3.2801 with p-value 0.194 indicating the series is normal distribution. This is an interesting property, and we defer this property to our further research.

Table 4

The basic statistics for the daily series of Nikkei 225 index over the period 7 October 1991 to 23 August 2005

	range	Log-range
Mean	1.659	0.3765
Median	1.449	0.3718
Maximum	8.929	2.189
Minimum	0.2906	-1.236
Standard deviation	0.9187	0.5077
Skewness	1.917	0.04863
Kurtosis	9.208	3.117
Jarque-Bera	7580.7	3.2801
Probability	0.000	0.194

We use Augmented Dickey-Fuller test to check whether the series is stationary or not. We use Schwarz Information Criterion (SIC),and set the maximum lag length 29. The equation we want to test includes intercept term. The results are display in table 5.

Table 5

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.49829	0.0000
Test critical values	1% level	-3.43208	
	5% level	-2.86219	
	10% level	-2.56716	

The t-statistics of the unit-root test is less than the critical value under 1% significance level. So the series is stationary and we can apply the series in the following empirical analysis.

Table 6 displays the estimation results for Nikkei 225 index series. For the series, we estimate a CARR model, a STCARR model with K=1 and a STCARR model with K=2. It is often found that CARR(1,1) specification is sufficient for a large class of financial data, and we estimate CARR(1,1) model to the series. The significance level with p-values correspond to w, α_1, β_1 are less than 0.001%, and the sum of α_1 and β_1 is less than 1. So the CARR(1,1) model can fit the series sufficiently well.

Table 6: Estimation results of CARR and STCARR models

$$R_t = \lambda_t \varepsilon_t$$

CARR(1,1): $\lambda_t = w + \alpha_1 R_{t-1} + \beta_1 \lambda_{t-1}$

STCARR(1,1): $\lambda_t = \sigma^2(1-\alpha_1-\beta_1) + \alpha_1 R_{t-1} + \beta_1 \lambda_{t-1} + \alpha_1^* R_{t-1} F(\ln R_{t-1}; \gamma, \mathbf{c})$

where $F(\ln R_{t-1}; \gamma, \mathbf{c}) = (1+\exp\{-\gamma(\ln R_{t-1} - c_1)\})^{-1}$

or $F(\ln R_{t-1}; \gamma, \mathbf{c}) = (1+\exp\{-\gamma(\ln R_{t-1} - c_1)(\ln R_{t-1} - c_2)\})^{-1}$

and σ^2 is unconditional variance

	CARR	STCARR(1,1) with K=1	STCARR(1,1) with K=2
w	0.02584(0.0000)		
α_1	0.1445(0.0000)	0.1607(0.0000)	0.1155(0.0000)
β_1	0.8400(0.0000)	0.8298(0.0000)	0.8251(0.0000)
γ		7.03(0.1218)	31
c_1		1.865(0.0000)	0.5394(0.0000)
c_2			1.417(0.0000)
α_1^*		-0.3299(0.0000)	-0.1918(0.0000)
LogL	-5635.3	-5328.9	-5596.9
LR-stat		612.8(0.000)	76.8(0.000)
LM-stat		4.522(0.033)	6.491(0.039)

Note:The figures in parentheses are p-values

It is difficult to obtain a very accurate estimate of the parameter γ, especially when γ is large. For the reason that large changes in the parameter γ have only minor effect on the value of the transition function, high accuracy in estimating γ is not necessary. When K=1, the staring values for the nonlinear optimization can be achieved by a two-dimensional grid search over γ and c_1. We rewrite the transition function using the method suggested in Dijk, Teräsvirta and Franses (2002) as follows.

$$F(\ln R_{t-1}; \gamma, \mathbf{c}) = (1+\exp\{-\frac{\gamma}{\hat{\sigma}}(\ln R_{t-1} - c_1)\})^{-1} \tag{18}$$

where $\hat{\sigma}$ is the sample standard deviation of $\ln R_{t-1}$.This modification makes γ approximately scale-free and helps us to choose a useful set of grid values for the parameter γ. The useful set of grid values for the parameter c_1 may be chosen as sample percentiles of the transition variable $\ln R_{t-1}$. This guarantees that the values of the transition function cover enough sample variation for each choice of γ and c_1.

It can be found in table 6 that K=1 own the smallest p-value for all the parameters except the parameter γ, whose p-value is not significant under 5% level. The estimate of c_1 is inside the range of $\ln R_{t-1}$. The low p-value of log likelihood ratio statistics indicate the significant improvements of fitness in STCARR(1,1) model with K=1 upon the associated CARR(1,1) model for the series. With some minor modification to the corollary 2 in appendix B, we derive the LM-version statistic to test the CARR(1,1) model against STCARR(1,1) model with K=1. The proof is very similar to the corollary 2 in appendix B and is omitted. The p-value of the LM statistic is under the 5% significant level which also indicate the evident improvement of fitness in STCARR(1,1) model with K=1 upon the CARR(1,l).

The estimated transition function $F(\ln R_{t-1};\gamma,\mathbf{c})$ with K=1 is plotted in Figure 2. It is termed the lower regime when $F(\ln R_{t-1};\gamma,\mathbf{c})$ =0 and the upper regime when $F(\ln R_{t-1};\gamma,\mathbf{c})$ reached its maximum value. During the transition period, $F(\ln R_{t-1};\gamma,\mathbf{c})$ shifts gradually from 0 to its maximum value. The negative sign of coefficient for α_1^* reveals the short-run impacts of unexpected innovations on volatilities in the lower regime is larger than that in the upper regime.

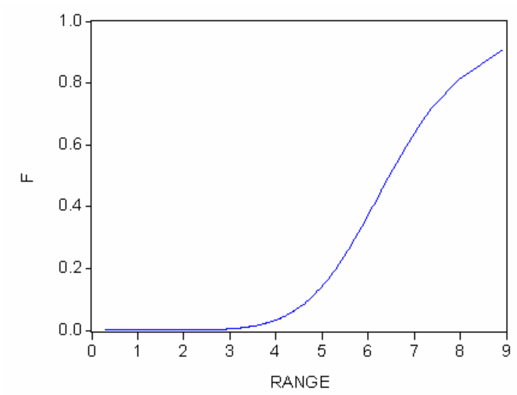

Figure 2: The Smooth Transition Function, when K=1

However we cannot work out the estimation result of the CARR(1,1) model with K=2 using the method mentioned above and we need some modifications. We think this is just because that the specification when K=2 is more complicated than the specification when K=1. Note that when the parameter γ in the transition function is known and fixed, the remained parameters in the CARR(1,1) model with K=2 are easily estimated using the similar grid search mentioned above. So the estimation problem can be reduced into maximizing the log likelihood function with respect to the parameter γ only.

It can be found in table 6 that when K=2 all the parameters except γ are significant under 1% level. The estimate of c_1 and c_2 are inside the range of $\ln R_{t-1}$. The low p-value of log likelihood ratio statistics reveals STCARR(1,1) model with K=2 outperforms the associated CARR(1,1) model for fitting the series. Using the LM-version test statistic derived in appendix B, we test the CARR(1,1) model against STCARR(1,1) model with K=2. The p-value of the LM statistic is under the 5% significant level which also indicate the significant improvement of fitness in STCARR(1,1) model with K=2 upon the associated CARR(1,l).

The estimated transition function $F(\ln R_{t-1};\gamma,\mathbf{c})$ with K=2 is plotted in Figure 2. It is found that the graph of $F(\ln R_{t-1};\gamma,\mathbf{c})$ is U-shape. It is termed the lower regime when $F(\ln R_{t-1};\gamma,\mathbf{c})$ =0 and the upper regime when $F(\ln R_{t-1};\gamma,\mathbf{c})$ =1. During the transition period, $F(\ln R_{t-1};\gamma,\mathbf{c})$ shifts gradually from 1 to 0 and back to 1. The negative sign of coefficient for α_1^* reveals the short-run impacts of unexpected innovations on volatilities in the lower regime is larger than that in the upper regime.

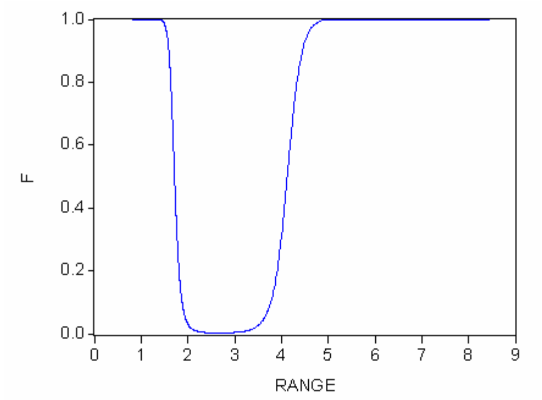

Figure 3: The Smooth Transition Function, when K=2

6. Conclusions

Since CARR model provides a simple and efficient way to model the volatility dynamics. For the reason that smooth transition models may be more realistic representation of aggregate behaviors in economies, we consider a new model named smooth transition CARR model which allows smooth transition mechanism to deal with the changes in regime in CARR specification. We have demonstrated empirically that smooth transition CARR model can improve fitness for financial time series upon the associated CARR model. Generalizations of the smooth transition CARR model will be meaningful subject of future research, for example, the generalization of the univariate to a multivariate framework, models simultaneous treating the return and the range, asymmetric volatility models and replacing the standard range with robust measures of range for the reason that range is sensitive to outliers.

Application of smooth transition CARR model to other asset prices such as exchange rate data will be meaningful. We will estimate smooth transition CARR models to stock index and exchange rate and compare the performances versus threshold CARR models.

It is mentioned above that the log-range series is normal distribution, this fact is consistent with some financial econometric literatures. So we can also choose the logarithmic range as a proxy for volatility and propose a new model named smooth transition LOGCARR model which is very similar to the smooth transition LOGACD model mentioned in Meitz and Teräsvirta (2006). In the spirit of smooth transition models, we can also allow the transition variable to be t in smooth transition CARR models.

References

Alizadeh, S., M.W.Brandt, and F.X.Diebold (2002). "Range-based estimation of stochastic volatility models. " *Journal of Finance*, 57:1047-1091.

Anderson, H. M., K. Nam, and F. Vahid (1999)."Asymmetric nonlinear smooth transition GARCH models." in Nonlinear Time Series Analysis of Economic and Financial Data, ed.by P. Rothman, pp. 191–207. Kluwer Academic Publishers, Boston.

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

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). "ARCH Modeling in Finance:A Review of the Theory and Empirical Evidence." *Journal of Econometrics* 52: 5–59.

Brandt,M. and C.Jones (2006)."Volatility Forecasting with Range-based EGARCH Models." *Journal of Business and Economic Statistics*,24:470-486.

Chou,R.Y.(2005)."Forecasting financial volatilities with extreme values: the Conditional AutoRegressive Range(CARR) Model. " *Journal of Money Credit and Banking*,37: 561-582.

Davies, R. B. (1977). "Hypothesis testing when a nuisance parameter is present only under the alternative." *Biometrika*, 64:247–254.

Davies, R. B. (1987). "Hypothesis testing when a nuisance parameter is present only under the alternative." *Biometrika*, 74:33–43.

Dueker,M.J.(1997)."Markov Switching in GARCH Processes and Mean Reverting Stock Market Volatility. ",*Journal of Business & Economic Statistics*,15:26-34.

Engle, R. F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation."*Econometrica*, 50: 987–1008.

Engle, R.F. (2000)."The econometrics of ultra-high-frequency data." *Econometrica,* 68:1–22.

Engle, R. F. (2002). "New Frontiers for Arch Models." *Journal of Applied Econometrics*,17:425–446.

Engle, R. F.,and J.R.Russell (1998). "Autoregressive Conditional Duration: A New Model for Irregular Spaced Transaction Data." *Econometrica*, 66:1127–1162.

Galant, A., C. Hsu, and G. Tauchen (1999). "Calculating volatility diffusions and extracting integrated volatility." *Review of Economics and Statistics*, 81:617-631.

Garman,M., and M. Klass (1980). "On the estimatin of security price volatilities from historical data," *Journal of Business,* 53:67-78.

Gonzalez-Rivera,G.(1998)."Smooth transition GARCH models." *Studies in Nonlinear*

Dynamics and Econometrics, 3:61–78.

Granger, C. W. J., and T. Teräsvirta (1993). “Modelling Nonlinear Economic Relationships.” New York: Oxford University Press.

Hagerud, G. E. (1997). “A new non-linear GARCH model.” PhD Dissertation, Stockholm School of Economics.

Hamilton,J.D., and P.Susmel(1994).“Autoregressive conditional heteroskedasticity and changes in regime.”*Journal of Econometrics,* 64:307-333.

Hansen, B. E. (1996). “Inference when a nuisance parameter is not identified under the null hypothesis.” *Econometrica*, 64:413–430.

Lin, C.-F.J.,and T. Teräsvirta(1994). “Testing the constancy of regression parameters against continuous structural change.”*Journal of Econometrics,* 62: 211-228.

Lukkonen, R., P. Saikkonen, and T. Teräsvirta (1988). “Testing linearity against smooth transition autoregressive models.”*Biometrika*, 75: 491–499.

Lundbergh, S., and T. Teräsvirta (1998). “Modelling economic high-frequency time series with STAR-STGARCH models.” working paper series in Economics and Finance No.291,Stockholm School of Economics.

Lundbergh, S., and T. Teräsvirta (2002). “Evaluating GARCH Models.” *Journal of Econometrics* 110: 417-435.

Meitz, M.,and T. Teräsvirta (2006). “Evaluating Models of autoregressive conditional duration.”*Journal of Business and Economics Statistics*,24:104-124.

Parkinson, M. (1980). “The Extreme Value Method for Estimating the Variance of the Rate of Return.” *Journal of Business*,53:61–65.

Rogers, C. and S. Satchell(1991).“Estimating variance from high, low and closing prices,” *Annals of Applied Probability* 1(4), 504-512.

Teräsvirta, T. (1994).“Specification, estimation and evaluation of smooth transition autoregressive models.”*Journal of the American Statistical Association,* 89: 208–218.

Van Dijk, D.,T.Teräsvirta and P.H.Franses (2002).“Smooth transition autoregressive models - a survey of recent developments”. *Econometric Reviews*,21:1-47

Yang, D. and Q. Zhang (2000). “Drift independent volatility estimation based on high, low, open, and close prices,” *Journal of Business*, 73:477-491.

Zhang, M.Y.,J. R. Russell, and R. S. Tsay (2001).“A nonlinear autoregressive conditional duration model with applications to financial transaction data.” *Journal of Econometrics,* 104:179–207.

Appendix A

Theorem1（Testing CARR models against general additive alternatives）

Consider the following model, with all the regularity conditions applied.

$$\begin{aligned} R_t &= (\lambda_t + \varphi_t)\varepsilon_t \\ \lambda_t &= \lambda_t(R_{t-1},...,R_1;\theta_1) \\ \varphi_t &= \varphi_t(R_{t-1},...,R_1;\theta_1,\theta_2) \\ \varepsilon_t &\sim i.i.d.\exp(1) \end{aligned} \qquad \text{(a.1)}$$

Assume that under $H_0:\theta_2=\theta_2^0$,the function $\varphi_t(R_{t-1},...,R_1;\theta_1,\theta_2^0)\equiv 0$ .let $\mathbf{a_t}(\theta_1)=\frac{1}{\lambda_t(\theta_1)}\frac{\partial\lambda_t(\theta_1)}{\partial\theta_1}$, $\mathbf{b_t}(\theta_1,\theta_2)=\frac{1}{\lambda_t(\theta_1)}\frac{\partial\varphi_t(\theta_1,\theta_2)}{\partial\theta_2}$, $c_t=\frac{R_t}{\lambda_t(\theta_1)}-1$ then, under the null hypothesis $H_0:\theta_2=\theta_2^0$,the LM-version statistic

$$LM=\{\sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t}^{'}\}\{\sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'}-(\sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'})(\sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'})^{-1}(\sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'})\}^{-1}\{\sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t}\} \qquad \text{(a.2)}$$

has an asymptotic χ^2 distribution with $\dim\theta_2$ degrees of freedom.

Proof:

Letting $\theta=(\theta_1^{'},\theta_2^{'})^{'}$ be the parameter vector and the conditional quasi log-likelihood function for observation R_t is defined as

$$l_t(\theta)=-\frac{R_t}{\lambda_t+\varphi_t}-\ln(\lambda_t+\varphi_t) \qquad \text{(a.3)}$$

The partial derivatives of (a.3) with respect to θ_1 and θ_2 are

$$\begin{aligned} \frac{\partial l_t(\theta)}{\partial\theta_1} &= \frac{1}{\lambda_t+\varphi_t}(\frac{\partial\lambda_t}{\partial\theta_1}+\frac{\partial\varphi_t}{\partial\theta_1})(\frac{R_t}{\lambda_t+\varphi_t}-1) \\ \frac{\partial l_t(\theta)}{\partial\theta_2} &= \frac{1}{\lambda_t+\varphi_t}\frac{\partial\varphi_t}{\partial\theta_2}(\frac{R_t}{\lambda_t+\varphi_t}-1) \end{aligned} \qquad \text{（a.4）}$$

Let $\theta^0=(\theta_1^{0'},\theta_2^{0'})^{'}$ be the true parameter vector under the null hypothesis, the score for observation R_t calculated at the true parameter values is

$$\frac{\partial l_t(\theta^0)}{\partial\theta}=\begin{bmatrix}\frac{1}{\lambda_t(\theta_1^0)+\varphi_t(\theta_1^0,\theta_2^0)}(\frac{\partial\lambda_t(\theta_1^0)}{\partial\theta_1}+\frac{\partial\varphi_t(\theta_1^0,\theta_2^0)}{\partial\theta_1})(\frac{R_t}{\lambda_t(\theta_1^0)+\varphi_t(\theta_1^0,\theta_2^0)}-1)\\ \frac{1}{\lambda_t(\theta_1^0)+\varphi_t(\theta_1^0,\theta_2^0)}\frac{\partial\varphi_t(\theta_1^0,\theta_2^0)}{\partial\theta_2}(\frac{R_t}{\lambda_t(\theta_1^0)+\varphi_t(\theta_1^0,\theta_2^0)}-1)\end{bmatrix}$$
$$=\left(\frac{R_t}{\lambda_t(\theta_1^0)}-1\right)\begin{bmatrix}\frac{1}{\lambda_t(\theta_1^0)}\frac{\partial\lambda_t(\theta_1^0)}{\partial\theta_1}\\ \frac{1}{\lambda_t(\theta_1^0)}\frac{\partial\varphi_t(\theta_1^0,\theta_2^0)}{\partial\theta_2}\end{bmatrix}=c_t(\theta_1^0)\begin{bmatrix}\mathbf{a_t}(\theta_1^0)\\ \mathbf{b_t}(\theta_1^0,\theta_2^0)\end{bmatrix}=c_t^0\begin{bmatrix}\mathbf{a}_t^0\\ \mathbf{b}_t^0\end{bmatrix} \tag{a.5}$$

Moreover, let $\hat{\theta}=(\hat{\theta}_1^{'},\theta_2^{0'})^{'}$ be the vector of maximum likelihood estimates under the null. Then the score evaluated at the ML estimates equals (note that the upper block of the score is now just a vector of zeros)

$$\frac{\partial l(\hat{\theta})}{\partial\theta}=\sum_{t=1}^{T}\frac{\partial l_t(\hat{\theta})}{\partial\theta}=\begin{bmatrix}\mathbf{0}\\ \sum_{t=1}^{T}c_t(\hat{\theta}_1)\mathbf{b_t}(\hat{\theta}_1,\theta_2^0)\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t}\end{bmatrix} \tag{a.6}$$

Using the condition (v) of Theorem 1 in Engle(2000).

$$T^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta}=T^{-1/2}\sum_{t=1}^{T}\frac{\partial l_t(\theta^0)}{\partial\theta}\xrightarrow{d}N\left(0,T^{-1}\sum_{t=1}^{T}E\left[\frac{\partial l_t(\theta^0)}{\partial\theta}\frac{\partial l_t(\theta^0)}{\partial\theta^{'}}\right]\right) \tag{a.7}$$

The expectation in this expression can be written as

$$E\left[\frac{\partial l_t(\theta^0)}{\partial\theta}\frac{\partial l_t(\theta^0)}{\partial\theta^{'}}\right]=E\left[(c_t^0)^2\begin{pmatrix}\mathbf{a_t^0a_t^{0'}} & \mathbf{a_t^0b_t^{0'}}\\ \mathbf{b_t^0a_t^{0'}} & \mathbf{b_t^0b_t^{0'}}\end{pmatrix}\right]=E\left[(c_t^0)^2\right]*E\left[\begin{pmatrix}\mathbf{a_t^0a_t^{0'}} & \mathbf{a_t^0b_t^{0'}}\\ \mathbf{b_t^0a_t^{0'}} & \mathbf{b_t^0b_t^{0'}}\end{pmatrix}\right] \tag{a.8}$$

since $c_t^0=\frac{R_t}{\lambda_t(\theta_1^0)}-1=\varepsilon_t-1$ is independent with the other terms, which are measurable with respect to $\mathbb{F}_{t-1}$. Furthermore, $E(c_t^0)^2=E(\frac{R_t}{\lambda_t(\theta_1^0)}-1)^2=Var(\varepsilon_t)=1$, so that

$$T^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta}\xrightarrow{d}N\left(0,T^{-1}\sum_{t=1}^{T}E\left[\begin{pmatrix}\mathbf{a_t^0a_t^{0'}} & \mathbf{a_t^0b_t^{0'}}\\ \mathbf{b_t^0a_t^{0'}} & \mathbf{b_t^0b_t^{0'}}\end{pmatrix}\right]\right) \tag{a.9}$$

This implies that the quadratic form

$$\left\{T^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta^{'}}\right\}\left\{T^{-1}\sum_{t=1}^{T}E\left[\begin{pmatrix}\mathbf{a_t^0a_t^{0'}} & \mathbf{a_t^0b_t^{0'}}\\ \mathbf{b_t^0a_t^{0'}} & \mathbf{b_t^0b_t^{0'}}\end{pmatrix}\right]\right\}^{-1}\left\{T^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta}\right\} \tag{a.10}$$

has an asymptotic χ^2 distribution with dim θ_2 degrees of freedom.

Since $\hat{\theta}$ is a consistent estimator of θ^0 (Theorem 1, Engle (2000)) and

$$T^{-1}\sum_{t=1}^{T} E\left[\begin{pmatrix} \mathbf{a_t}(\hat{\theta}_1)\mathbf{a_t}^{'}(\hat{\theta}_1) & \mathbf{a_t}(\hat{\theta}_1)\mathbf{b_t}^{'}(\hat{\theta}_1,\theta_2^0) \\ \mathbf{b_t}(\hat{\theta}_1,\theta_2^0)\mathbf{a_t}^{'}(\hat{\theta}_1) & \mathbf{b_t}(\hat{\theta}_1,\theta_2^0)\mathbf{b_t}^{'}(\hat{\theta}_1,\theta_2^0) \end{pmatrix}\right] = T^{-1}\sum_{t=1}^{T}\begin{bmatrix} \hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \\ \hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \end{bmatrix} \quad \text{(a.11)}$$

is a consistent estimator of

$$T^{-1}\sum_{t=1}^{T} E\left[\begin{pmatrix} \mathbf{a_t^0 a_t^{0'}} & \mathbf{a_t^0 b_t^{0'}} \\ \mathbf{b_t^0 a_t^{0'}} & \mathbf{b_t^0 b_t^{0'}} \end{pmatrix}\right] \quad \text{(a.12)}$$

(since the expressions are functions of $\hat{\theta}$ and θ^0, respectively), the LM statistic

$$\left\{T^{-1/2}\frac{\partial l(\hat{\theta})}{\partial \theta^{'}}\right\}\left\{T^{-1}\sum_{t=1}^{T}\begin{bmatrix} \hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \\ \hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \end{bmatrix}\right\}^{-1}\left\{T^{-1/2}\frac{\partial l(\hat{\theta})}{\partial \theta}\right\}$$

$$= \begin{bmatrix} \mathbf{0} \\ \sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t} \end{bmatrix}^{'}\left\{\begin{bmatrix} \sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \\ \sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'} & \sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} \end{bmatrix}\right\}^{-1}\begin{bmatrix} \mathbf{0} \\ \sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t} \end{bmatrix} \quad \text{(a.13)}$$

$$= \{\sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t}^{'}\}\{\sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'} - (\sum_{t=1}^{T}\hat{\mathbf{b}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'})(\sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{a}}_\mathbf{t}^{'})^{-1}(\sum_{t=1}^{T}\hat{\mathbf{a}}_\mathbf{t}\hat{\mathbf{b}}_\mathbf{t}^{'})\}^{-1}\{\sum_{t=1}^{T}\hat{c}_t\hat{\mathbf{b}}_\mathbf{t}\}$$

also has an asymptotic χ^2 distribution with $\dim\theta_2$ degrees of freedom under the null hypothesis.

In application the LM statistic in Theorem 1 can be easily carried out using an auxiliary least squares regression on particular transformed variables. This can be done as the following Theorem 2.

Theorem 2

(1)Obtain the quasi maximum likelihood estimator of θ_1 under the null hypothesis,and compute $\hat{\mathbf{a}}_\mathbf{t}^{'}(\widehat{\theta}_1) = \frac{1}{\lambda_t(\widehat{\theta}_1)}\frac{\partial\lambda_t(\widehat{\theta}_1)}{\partial\theta_1^{'}}$, $\hat{\mathbf{b}}_\mathbf{t}^{'}(\widehat{\theta}_1,\theta_2^0) = \frac{1}{\lambda_t(\widehat{\theta}_1)}\frac{\partial\varphi_t(\widehat{\theta}_1,\theta_2^0)}{\partial\theta_2^{'}}$,

$\widehat{c}_t = \frac{R_t}{\lambda_t(\widehat{\theta}_1)} - 1, t = 1,...,T$ and $SSR_0 = \sum_{t=1}^{T}\hat{c}_t^2$.

(2)Regress $\widehat{c}_t$ on $\hat{\mathbf{a}}_\mathbf{t}^{'}$ and $\hat{\mathbf{b}}_\mathbf{t}^{'}$, $t = 1,...,T$,and calculate SSR_1.

Then,under the null hypothesis, the test statistic (a.2) can be achieved by

LM= $T(SSR_0 - SSR_1)/SSR_0$ (a.14)

which has an asymptotic χ^2 distribution with $\dim\theta_2$ degrees of freedom

the F-version test statistic of the statistic (a.14) is

$$F = \frac{(SSR_0 - SSR_1)/\dim\theta_2}{SSR_1/(T-p-q-1-\dim\theta_2)} \quad \text{(a.15)}$$

which has an asymptotic F distribution with ($\dim\theta_2, T-p-q-1-\dim\theta_2$) degrees of freedom

Appendix B

Corollary 2

Here we use corollary 1 to test the CARR(1,1) against smooth transition CARR(1,1). Consider the model as follows

$$\begin{aligned}
R_t &= (\lambda_t + \varphi_t)\varepsilon_t \\
\lambda_t + \varphi_t &= w + \alpha_1 R_{t-1} + \beta_1(\lambda_{t-1} + \varphi_{t-1}) + e_{11}R_{t-1}(\ln R_{t-1}) + e_{12}R_{t-1}(\ln R_{t-1})^2 \\
\varphi_t &= \beta_1\varphi_{t-1} + e_{11}R_{t-1}(\ln R_{t-1}) + e_{12}R_{t-1}(\ln R_{t-1})^2 \\
\varepsilon_t &\sim i.i.d.\exp(1)
\end{aligned} \quad \text{(b.1)}$$

denote $\theta_1 = (w, \alpha_1, \beta_1)'$, $\theta_2 = (e_{11}, e_{12})'$ and $J = \left(R_{t-1}(\ln R_{t-1}), R_{t-1}(\ln R_{t-1})^2\right)'$. under the null hypothesis $H_0 : \theta_2 = 0$. where

$$\begin{aligned}
\hat{\mathbf{a}}_t &= \frac{1}{\hat\lambda_t}\frac{\partial\hat\lambda_t}{\partial\theta_1} = \hat\lambda_t^{-1} \times (1, R_{t-1}, \hat\lambda_{t-1})' + \hat\lambda_t^{-1}\hat\beta_1\frac{\partial\hat\lambda_{t-1}}{\partial\theta_1} \\
&= \frac{1}{\hat\lambda_t}\left[\sum_{i=1}^{t-1}\hat\beta_1^{i-1}, \sum_{i=1}^{t-1}\hat\beta_1^{i-1}R_{t-i}, \sum_{i=1}^{t-1}\hat\beta_1^{i-1}\hat\lambda_{t-i}\right]' \\
\hat{\mathbf{b}}_t &= \frac{1}{\hat\lambda_t}\frac{\partial\hat\varphi_t}{\partial\theta_2} = \hat\lambda_t^{-1} \times J + \hat\lambda_t^{-1}\hat\beta_1\frac{\partial\hat\varphi_{t-1}}{\partial\theta_2} \\
&= \frac{1}{\hat\lambda_t}\left[\sum_{i=1}^{t-1}\hat\beta_1^{i-1}R_{t-i}\ln(R_{t-i}), \sum_{i=1}^{t-1}\hat\beta_1^{i-1}R_{t-i}\ln(R_{t-i})^2\right] \\
\hat c_t &= \frac{R_t}{\lambda_t(\hat\theta_1)} - 1
\end{aligned} \quad \text{(b.2)}$$

Compute $SSR_0 = \sum_{t=1}^{T} \hat{c}_t^2$. Regress $\hat{c}_t$ on $\hat{\mathbf{a}}_\mathbf{t}^{'}$ and $\hat{\mathbf{b}}_\mathbf{t}^{'}$, $t = 1,...,T$,and compute SSR_1 .Then under the null hypothesis, the LM-version test statistic

$$\text{LM}=T(SSR_0 - SSR_1)/SSR_0 \quad \text{(b.3)}$$

has an asymptotic χ^2 distribution with two degrees of freedom

the F-version test statistic of the statistic (b.3) can be computed by

$$F = \frac{(SSR_0 - SSR_1)/2}{SSR_1/(T-1-1-1-2)} \quad \text{(b.4)}$$

which has an asymptotic F distribution with $(2, T-5)$ degrees of freedom