WORKING PAPER NO: 499

Retesting the estimation of a utility-based asset pricing model using normal mixture GARCH (1, 1)

Rohit Gupta
FPM Student
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
rohit.gupta09@iimb.ernet.in

Vinu C T
FPM Student
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
vinu.ct10@iimb.ernet.in

Year of Publication – December 2015

Abstract

The main purpose of this paper is to derive the process of estimating dynamic RRA with the maximum likelihood and a Bayesian method having a weakly informative prior density while assuming that the log excess returns on the market are distributed as normal mixture, GARCH(1,1), Mixture GARCH (1, 1). Simulation analysis has been used to compare MLE and Bayesian estimates. Empirical results using GARCH model are presented using market rates of returns and risk-free rates over the period 1941 to 2010.

Keywords: Bayesian, risk aversion, normal mixture, MLE, simulation

1 Introduction

The theory[1] and the distribution of unconditional relative risk aversion (RRA) estimates in utility-based asset pricing model by assuming normality for the log excess returns has been developed in past few decades. While the normality assumption is not always appropriate for some security returns. Generalized method of moments (GMM) has been proposed[2] to estimate unconditional RRA. However, RRA estimated by GMM is not statistically efficient with finite samples.

An individual's preference for risky alternatives is influenced by the strength of preference he feels for the consequences and his attitude toward risk taking. The power utility function[3], one among utility-based asset pricing models, defined over consumption states with coefficient of relative risk aversion (RRA) and rate of time preference is the most commonly utilized preference specifications.

Furthermore, because the unknown parameter of isoelastic utility function is RRA, Brown and Gibbons (1985) and Alonso et al. (1990) have argued that using this utility function can provide precious information for several reasons. [4] Moreover, few of the used Studies[5] used aggregate consumption data for empirical analyses, which are measured with error and are time-aggregated, will have serious consequence for asset pricing relationships.[6] Later, few of the studies on utility-based asset pricing models replaced aggregate consumption with the return on some proxy for the market portfolio; this substitution is also employed in Wu & Lee (2007) to avoid measurement problems with consumption.[7]

Analyzing the distribution of asset returns data has also been an important research area in financial economics. Accordingly, utility-based models of the asset pricing are of particular interest while the distribution of returns can be suitably determined and explained. When the excess return on the market portfolio is distributed as a lognormal distribution, different methods for estimating unconditional relative

[1] A simple econometric approach for utility based asset pricing model. Journal of Finance 40, 359–381], Karson et al. [Karson, M., Cheng, D., Lee, C. F., 1995. Sampling distribution of the relative risk aversion estimator: theory and applications. Review ofQuantitative Finance and Accounting 5, 43–54], and Lee et al. [Lee, C.F., Lee, J.C., Ni, H.F., Wu, C.C., 2004. On a simple econometric approach for utility-based asset pricing model. Review of Quantitative Finance and Accounting 22, 331–344]
[2] Brown and Gibbons [Brown, D.P., Gibbons, M.R., 1985. A simple econometric approach for utility-based asset pricing model. Journal of Finance 40, 359–381]
[3] Hansen and Singleton (1982, 1983) postulated that the pricing kernel is a power function of aggregate U.S. consumption and estimated the parameters of the power pricing kernel. Hansen and Jagannathan (1991) derived bounds of the consumption-based pricing kernel from the mean and standard deviation of the market portfolio excess returns.
[4] First, since some theoretical results in finance rely on log utility function (i.e., β=1) (see, Hakanson (1970), Kraus and Litzenberger (1975), Rubinstein (1977), Cox, Ingersoll, and Ross (1985)), appropriate judgment on these results have to be made when RRA is significantly different from one. Second, when we are confronted with the demand for risky assets and the savings decisions, the demand for risky assets depends on the magnitude of RRA (see, Rothschild and Stiglitz, 1971). Third, there are many research papers dealing with the issue whether stock prices have excessive volatility relative to the degree of aggregate risk aversion (see, Grossman and Shiller, 1981).
[5] Lucas (1978), Grossman and Shiller (1981), Duan and Singleton (1986)
[6] Campbell (1993) and Rosenberg and Engle (2002)
[7] Brown and Gibbon (1985), Bansal and Viswanatham (1993), Campbell (1993), and Rosenberg and Engle (2002)

risk aversion (RRA), β, have been proposed[8] and the exact sampling and Bayesian estimators of RRA has been derived.

However, it is well known that the normal distribution may not be adequate[9] for the log asset returns. Therefore, Brown and Gibbons (1985) dropped the distributional assumption and recommended using generalized method of moments (GMM)[10] to estimate unconditional RRA. But, estimating RRA based on unconditional and fixed moments will not capture the phenomenon of structural changes. Even though a variety of evaluation methods have been proposed and implemented to RRA estimator to date, they have mostly depended on GMM or the classical statistical framework.

Pastor and Stambaugh (2000) assumed normality of returns and presented the Bayesian set-up, which factored uncertainties in both parameter estimation and model mispricing into investor's decision making. Although many appropriate distributions have been proposed to analyze asset returns, the RRA estimator does not always have a solution by assuming asset returns distributed as any distribution.

Thus, Wu and Lee (2007) extended Brown and Gibbons (1985) and used the NM (K)-GARCH model, i.e. the model where errors have K-component normal mixture distribution with generalized autoregressive conditional Heteroscedasticity (GARCH) variance process, to obtain more efficient dynamic RRA estimator and its distribution. The primary purpose of their paper was to derive the process of estimating dynamic RRA while assuming that the log excess returns on the market are distributed as NM (K) - GARCH (1, 1). On the part of the parameter estimation, they not only presented the classical maximum likelihood but also recommended a Bayesian method, which combines an investor's prior belief about the accuracy of the pricing model and the information in the data, with a weakly informative prior density.

Along with the model discussed in the Wu and Lee (2007) we have tried two simple versions of the model. The estimation part of NM-GARCH model was time taking and made us to think about simpler models. Model 1 as NM-GARCH and the other two models are Mixture of Two normal (Model 2) and GARCH model (Model 3). Model 2 and Model 3 are special cases of Model 1. Simulation of Model 2 and Model 3 is used for comparison of Bayesian and MLE estimation. An empirical study using model 3 is given in the later section.

---

[8] Brown and Gibbons (1985), Karson et al. (1995), and Lee et al. (2004)

[9] The empirical findings seem to indicate that the unconditional and conditional distributions of log asset returns are not symmetric and have fat tails relative to the normal distribution.

[10] GMM estimates and their standard errors are consistent even though residuals are heteroscedastic. However, the GMM can be applied only to large samples. In most cases GMM estimates are asymptotically efficient, but they are hardly efficient at finite samples. In addition, we found that the unconditional RRA estimates vary a lot across different sub-periods in Brown and Gibbons (1985) and, Lee et al. (2004).

2 Literature Review

The well-known Euler condition for the dynamic consumption-portfolio problem faced by a representative individual under uncertainty is used to derive a relationship between relative risk aversion and the moments of security returns. This first-order necessary condition for optimality with a time additive, von Neumann-Morgenstern utility function is

$$E\left[\frac{1}{1+r}\frac{U'(\tilde{C}_t)}{U'(C_{t-1})}(1+\tilde{R}_{it}) \mid Z_{t-1}^*\right] = 1 \; \forall\, i = 1,2,\ldots,N \; \forall\, t = 1,2,\ldots.T \qquad (1)$$

Where

$U'(\tilde{C}_t)$ = Marginal utility in period t from consumption, $\tilde{C}_t$
r= Rate of pure time discount
$\tilde{R}_{it}$= Return on asset i in period t
Z_{t-1}^*= Information set available in the market in period t-1

This first-order condition reflects the loss of marginal utility of consumption today if one additional share of a security is purchased versus the gain in expected marginal utility tomorrow when the share is sold and the return is consumed. [11]

This equation holds for all time horizons.[12] Without additional assumptions, Equation (1) provides little guidance for empirical research. In the work that followed[13], power (or isoelastic) utility[14] is specified; that is,

$$U(C_t) = \frac{C_t^{1-\beta}-1}{1-\beta} \qquad (2)$$

Where β = - U"(C_t) C_t/U'(C_t), relative risk aversion. By assuming isoelastic utility, Equation (2) can be restated as

[11] Among others, Lucas [25] discusses this equation in the discrete time case, and Grossman and Shiller [13] provide a derivation in the continuous time setting. Hansen, Richard, and Singleton [20] have also emphasized the importance of Equation (1) for econometric analyses of asset pricing models.
[12] If agents make decisions daily, the equation is still relevant to the econometrician who has data sampled at monthly intervals; thus, temporal aggregation bias may be avoided.
[13] Brown and Michael R. Gibbons (1985)
[14] Power utility is a natural choice because of its desirable theoretical properties. As a member of the HARA class of utility functions, Rubinstein [33] has established its aggregation properties over individuals in the economy. Further, log utility is a special case of isoelastic utility as β approaches one. Since β is estimated, its distance from one can be determined.
Also Arrow [1971] has emphasized that absolute risk aversion should be decreasing, and power utility displays this characteristic.

$$E\left[\frac{1}{1+r}\left(\frac{\tilde{C}_t}{C_{t-1}}\right)^{\beta}(1+\tilde{R}_{it})\ \mid Z_{t-1}^{*}\right]=1$$

(3)

If appropriate measures of consumption were available, Equation (3) could be transformed to yield empirical implications. However, much of the early research on utility-based asset pricing models replaced aggregate consumption with the return on some proxy for the market portfolio; this substitution is also employed here to avoid measurement problems with consumption as well as to relate to these earlier studies. With additional assumptions Equation (3) can be rewritten as:

$$E\left[\frac{1}{1+r}(1-k)^{-\beta}(1-\tilde{R}_{mt})^{-\beta}(1+\tilde{R}_{it})\ \mid Z_{t-1}^{*}\right]=1$$

(4)

Where

$R_{mt} =$ Return on the market portfolio

k = proportion of wealth consumed in every period (i.e., if C is consumption and W is wealth, then C = kW)

Hakansson [1970] pointed out that consumption is a constant proportion of wealth if individuals with infinite horizons have isoelastic utility functions, and the distribution of real production opportunities is constant and characterized by constant stochastic returns to scale. Alternatively, in a pure exchange economy, Rubinstein [1974] demonstrates that if aggregate consumption growth rates are independently and identically distributed over time or if β equals one then (4) results with the return on the market index being independently and identically distributed as well.[15]

Substituting R_{mt} and R_{ft} (the return on the one period riskless bond) for R_{it} in (4) and equating the left-hand sides of the resulting relations, it is evident after a simple rearrangement that

$$E\left[\left(1+\tilde{R}_{mt}\right)^{1-\beta}\mid Z_{t-1}^{*}\right]=(1+r)(1-k)^{\beta}=\ E\left[\left(1+\tilde{R}_{mt}\right)^{-\beta}(1+R_{ft}\ )\ \mid Z_{t-1}^{*}\right]$$

$$E\left[\left(1+\tilde{R}_{mt}\right)^{1-\beta}\mid Z_{t-1}^{*}\right]-\ E\left[\ \left(1+\tilde{R}_{mt}\right)^{-\beta}\left(1+R_{ft}\ \right)\mid Z_{t-1}^{*}]\right]=0$$

(5)

Since a riskless asset has a non-stochastic return conditional on Z_{t-1}^{*}, R_{ft} can be brought outside the expectation operator. With a little algebraic manipulation,

[15] In fact, (4) is consistent with an equilibrium in any production economy in which aggregate consumption growth rates are independently and identically distributed.

$$E[(x_t - 1){x_t}^{-\beta} | Z^*_{t-1}] = 0$$

(6)

Where x_t = (1 + R_{mt})/ (l + R_{ft}), a discrete time "excess return" Equation (6) has implications not only for moments conditional on complete information but also conditional on coarser information. Assuming the relevant moments exist, the law of iterated expectations applied to (6) implies

$$E\left[(x_t - 1){x_t}^{-\beta}\right] = 0$$

(7)

Although individuals have changing conditional expectations through time, the econometrician, by relying on Eq. (3), can still construct a valid test of the theory based on unconditional and fixed moments. Hence, most previous literatures claimed that it is not necessary to specify a model for the conditional expectations or even the variables which affect these conditional expectations and only estimated the static RRA, called β instead of β_t in this paper, during some specific periods.

Rubinstein (1976) proposed estimating β by assuming that the excess return on the market has a lognormal distribution, i.e., log x_t ~N(μ,σ^2), and derived

$$\beta = \frac{\mu}{\sigma^2} + \frac{1}{2}$$

(8)

Brown and Gibbons (1985), a natural maximum likelihood estimator for β is

$$\hat{\beta} = \frac{\hat{\mu}}{\hat{\sigma}^2} + \frac{1}{2}$$

Where

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}\log x_t \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\log x_t - \overline{\log x_t}\right)^2$$

Using the asymptotic theory they also derived the variance of $\sqrt{T}\,\hat{\beta}$ as:

$$\text{Var}\,(\sqrt{T}\,\hat{\beta}) = \frac{2[E\,(\log x_t\,)]^2 + Var[\log x_t\,]}{Var[\log x_t\,]^2}$$

(9)

Alternatively, following Karson et al. (1995), the minimum variance unbiased (MVU) estimator of β is

$$\hat{\beta} = \frac{(T-3)+\hat{\mu}}{(T-1)+\hat{\sigma}^2} + \frac{1}{2}$$
(10)

Brown and Gibbons (1985) pointed out that estimate of β is inconsistent when the distribution of log excess returns is a departure from normality. In order to remedy this weakness, they made use of the generalized method of moments to estimate RRA. Therefore, RRA could be estimated by finding the value β^ which satisfies the equation,

$$f(\hat{\beta}) = \frac{1}{T}\sum_t (x_t - 1)x_t^{-\hat{\beta}}] = 0$$
(11)

Where $\hat{\beta}$ ∈(0, +∞), and the asymptotic variance is

$$\text{Var}\,(\sqrt{T}\,\hat{\beta}) = \frac{E\left\{\left[(x_t-1)x_t^{-\hat{\beta}}\right]^2\right\}}{\left[E\left\{(x_t-1)x_t^{-\hat{\beta}} log x_t\right\}\right]^2}$$
(12)

3 Methodology & Estimation

Bayesian and Maximum likelihood estimation of models discussed here. EM algorithm is used for the maximum likelihood estimation of finite mixture model. Bayesian estimation using posterior simulation techniques such as Markov chain Monte Carlo (MCMC) methods.

Owing to its greater flexibility, the normal mixture model has also been found superior for describing the distribution of asset returns. A mixture of normal distributions has been proposed as a suitable framework for capturing the skewness. [16]

The family of K-component normal mixtures is capable of exhibiting the skewness and excess kurtosis characteristics of economical and financial data. The advantages of normal mixtures are the valid distributional assumption and the ease in financial explanations. For example, the normal mixture formulation allows for a sensible interpretation of two or more heterogeneous groups of market participants. Consequently, bullish and bearish investors could behave differently. Moreover, the mixture can also be interpreted as representing trading days of different types. For instance, returns on the stock market on Mondays will follow the prevailing trend from the previous Friday, and this phenomenon is viewed as the Monday effect. Besides, there are also January effect, October effect, etc., and these phenomena may cause the stock's behavior in specific months to depart from the other trading months.

[16] Positive skewness in the individual returns of the stocks that make up the Dow Jones Index

Furthermore, some substantial political or economical policies may give rise to a component with relatively high or low variance and smaller weight.

Since Engle (1982) first found that volatility shows a clustering phenomenon, numerous academic literatures have proposed a variety of models to predict future volatilities and exhibit market volatilities which are more predictable than market returns. However, most of those papers assume that the conditional distributions of the asset returns are symmetric, such as normal or student-t models. Recently, few authors improved these drawbacks and proposed NM (K)-GARCH (p, q) models with an inter-dependent autoregressive evolution for the variance series to capture the non-zero conditional excess kurtosis and skewness. However, the results of Haas et al. (2004) exhibited that neither the dependence of component variances nor the inclusion of more than one lag in the conditional variance equations are significant. Therefore, Alexander and Lazar (2005) recommended using a simpler NM (K)-GARCH (1, 1) models to Exchange rate modeling. Throughout this paper, we estimate RRA by applying NM (K)-GARCH (1, 1) model to log market excess returns.[17]

The log market excess returns, Y_t, distributed as NM (K)-GARCH (1, 1), is specified as follows:

$$Y_t|\Im_{t-1} \sim NM(p_1, \dots, p_k; \mu_1, \dots \mu_k;\ \sigma_{1,t}^2, \dots, \sigma_{k,t}^2),$$

$$\sigma_{j,t}^2 = w_j + a_j\varepsilon_{t-1}^2 + a_j\sigma_{j,t-1}^2\ ,$$
(13)

Where,
$$\textstyle\sum_{j=1}^{k} p_j = 1\ , \sum_{j=1}^{k} p_j = 1, \sum_{j=1}^{k} p_j\mu_j = \mu\ , \varepsilon_{t-1} = y_{t-1} - \mu,$$

$$w_j > 0, a_j, b_j > 0, and, a_j + b_j < 1\ for\ j = 1, \dots, K.$$

The conditional density function of $y_t\ is\ f(y_t|\Im_{t-1}) = \sum_{j=1}^{K} p_j\phi_j(Y_t|\,\Im_{t-1})\ and\ \phi_j((y_t|\Im_{t-1})$ is a normal density function with mean μ_j and variance $\sigma_{j,t}^2$

For the mixture model, the number of required component densities is unknown and need to be empirically determined generally. Unfortunately, standard likelihood ratio test (LRT) theory breaks down in this framework. However, standard model selection criteria such as AIC and BIC are widely used to compare models with different numbers of components. For a model with d parameters, sample size T and log-likelihood ℓ, evaluated at the maximum likelihood estimator, AIC= $-2\ell+2d$ and BIC=$-2\ell+d \log T$. When $\log T > 2$, it can be found that the penalty term of BIC penalizes complex models more heavily than AIC. Hence, AIC tends to fit too many components, and BIC is more conservative in that it favors more parsimonious models. Furthermore, the literatures on mixtures provide some encouraging evidence in the

[17] Wu and Lee (2007)

context of unconditional model and suggest that BIC provides a reasonably good indication for the number of components.

3.1 Model 1 NM-GARCH

The author has tried GMM, MLE, and Bayesian estimation procedure for estimating the model parameters. The following section included the Bayesian estimation details.

3.1.1 Estimation of RRA

For the transformed version of Eq. (2), we assume that log x_t , the log excess returns on the market follows NM (K) ~GARCH (1, 1) model. So,

$$y_t|\Im_{t-1} = \log x_t|\Im_{t-1} \sim NM\big(p_{1,} \dots \dots, p_k; \mu_1, \dots \dots, \mu_k;\ \sigma^2_{1,t}, \dots \dots, \sigma^2_{k,t}\big),$$

And consequently we have

$$E[x_t^{-\beta t}|\Im_{t-1}] = \sum_{j=1}^{k} p_j\, e^{-\mu_j\beta_t + \frac{\sigma^2_{j,t}}{2}}$$
(14)

Substitution in Eq. (2) gives

$$g\,(\beta_t|\Im_{t-1}, \mu_j, \sigma^2_{j,t}, j=1,\dots,k) = \sum_{j=1}^{k} p_j\, e^{\mu_j(1-\beta_t) + \frac{\sigma^2_{j,t}(1-\beta_t)^2}{2}} - \sum_{j=1}^{k} p_j\, e^{-\mu_j\beta_t + \frac{\sigma^2_{j,t}\beta_t^2}{2}} = 0 \qquad (15)$$

Given that the excess returns on the market portfolio is NM (K)-GARCH (1, 1) model, Eq. (15) provides a relationship between dynamic RRA and the parameters of the NM (K)-GARCH (1, 1) model. Since log $x_t|Z_{t-1}$ has a normal mixture distribution, the MLE of RRA has no analytical form. However, because a continuous function of maximum likelihood estimators is also a maximum likelihood estimator (Zehna 1966), $\widehat{\beta}_t$ satisfying the following equation has all of the well-known properties of the maximum likelihood estimator,

$$g\,(\widehat{\beta}_t|\Im_{t-1}, \widehat{\mu}_j, \widehat{\sigma}^2_{j,t}, j=1,\dots,k) = \sum_{j=1}^{k} \widehat{p}_j\, e^{\widehat{\mu}_j(1-\widehat{\beta}_t) + \frac{\widehat{\sigma}^2_{j,t}(1-\widehat{\beta}_t)^2}{2}} - \sum_{j=1}^{k} \widehat{p}_j\, e^{-\widehat{\mu}_j\widehat{\beta}_t + \frac{\widehat{\sigma}^2_{j,t}\widehat{\beta}_t^2}{2}} = 0$$
(16)

The function g() is a strictly decreasing continuous function, so Eq. (16) has a unique solution. In addition, Fig. 1 illustrates the nonlinear search for β_t, when the other parameters are fixed.

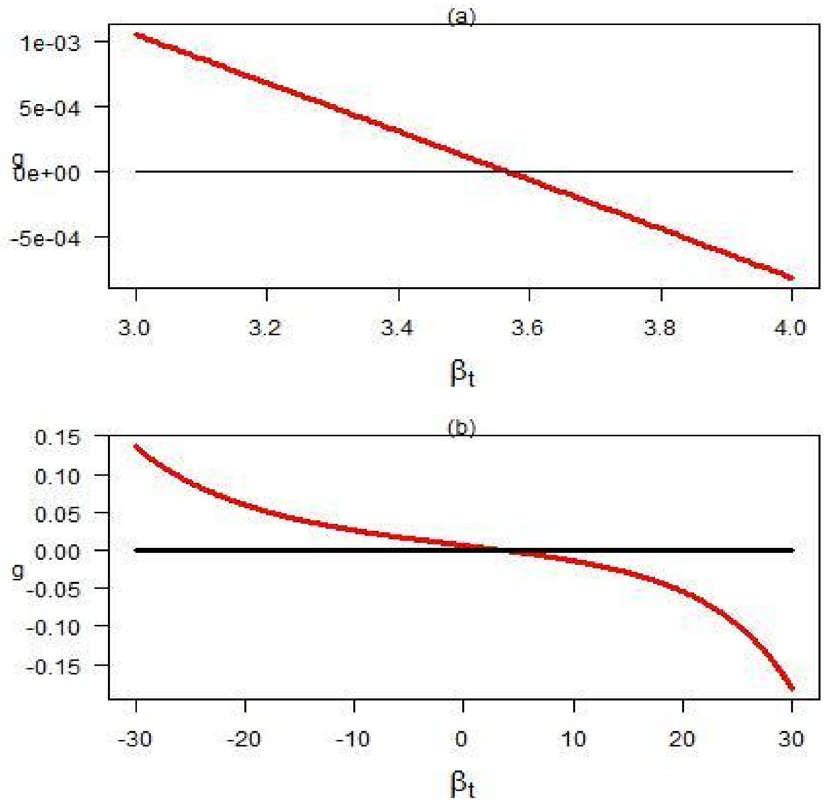


Fig. 1. Graphical determination of the value of RRA, β_t for which the equilibrium condition is satisfied, i.e. $g\left(\beta_t \middle| Z_{t-1}, \mu_j, \sigma_{j,t}^2, j = 1, \dots, K\right) = 0$. The exhibitions are based on K=2, (p_1, p_2) = (0.3, 0.7), $(\mu_1, \mu_2) = (-0.014, 0.014)$, and $\left(\sigma_{1,t}^2, \sigma_{2,t}^2\right) = (0.003, 0.001)$.

3.1.2 Bayesian Estimation of RRA

We have seen that estimation for RRA based on NM (K)-GARCH (1, 1) models is straightforward using the EM algorithm. Meanwhile, with the advent of inexpensive and high speed computers, estimation in Bayesian framework is now feasible using posterior simulation via Markov chain Monte Carlo (MCMC) methods. From past experiences, we would not expect inference about the parameters in Eq. (17)

$$L(\psi) = \prod_{t=2}^{T}\left[p_j \phi_j(y_t \Im_{t-1})\right]$$
(17)

to be highly sensitive to prior specification. In general, we may prefer non-informative priors to informative priors, if no prior information is available. Nevertheless, the main hindrance in normal mixture models is that improper non-informative priors will not yield proper posterior distributions (Diebolt and Robert, 1994; Roeder and Wasserman, 1997). Therefore, in this section, we choose a fixed number of components, K, according to BIC and refer to Richardson and Green (1997) to construct weakly informative priors for model parameters. As in Richardson and Green (1997), we assume that μ_j are drawn independently with normal priors,

$$\mu_j \sim N(\xi, \kappa^{-1})$$
(18)

For the variance processes, we assume priors between the θ= $(w_j, a_j, b_j,)$ are independent uniform distributions, which are given by

$$\prod(w, a, b) = \prod_{j=1}^{K} \prod(w_j, a_j, b_j) \propto \prod_{j=1}^{K} \quad I(w_j > 0, a_j \geq 0, b_j \geq 0, a_j + b_j < 1)$$
(19)

Where I (·) is the indicator function with I(S) =1 if the event S is true, otherwise I(S) =0. The prior on the weights p= $(p_1, p_2, \ldots\ldots, p_k)$ will always be taken as symmetric Dirichlet,

$$p \sim D(\delta, \delta, \ldots \delta,)$$
(20)

In order to give weakly informative priors for the model parameters, we introduce hyper-prior and hyper-parameter choices which correspond to making the minimal assumption on the data. Before determining the hyper-parameters, we comment briefly on the issue of labeling the components. The whole model is invariant with respect to permutation of the labels j=1, 2… K. For identifiability, Richardson and Green (1997) adopt a unique labeling in which the μ_j are in increasing numerical order. Hence, the joint prior distribution of the μ_j is K! Times the product of the individual normal densities, restricted to the set $\mu_1 < \mu_2 < \mu_3 < \mu_4 \ldots . < \mu_k$. Following Richardson and Green (1997) we take the N(ξ, κ−1) prior for μ_j to be rather flat over the range of data, by letting ξ equal to the mid-point of this range, and κ equal to a small multiple of 1/R2, where R is the length of the range. The complete hierarchical model is displayed in Fig. 2 as a directed acyclic graph (DAG). We follow the usual convention of graphical models that square boxes represent fixed or observed quantities and circles represent the unknowns.

In order to make inferences about model parameters, we need to integrate over high dimensional probability distribution, which could be very difficult. MCMC methods are very helpful for solving our problem. MCMC is essentially Monte Carlo integration using Markov chains. It draws samples from the required distribution by running a cleverly constructed Markov chain for a long time and then forms sample averages to approximate expectations. The Gibbs sampler and the Metropolis–Hastings (M–H) algorithm are well known among the several ways of constructing those chains. A great advantage of the Gibbs sampler and the M–H algorithm is the ease of implementation which makes heavy use of the modern computational capabilities. Excellent tutorials on the methodology have been provided by Casella and George (1992), Chib and Greenberg (1995), Gilk, Richardson and Spiegel halter (1996). The MCMC methods are used to make inferences in this section.

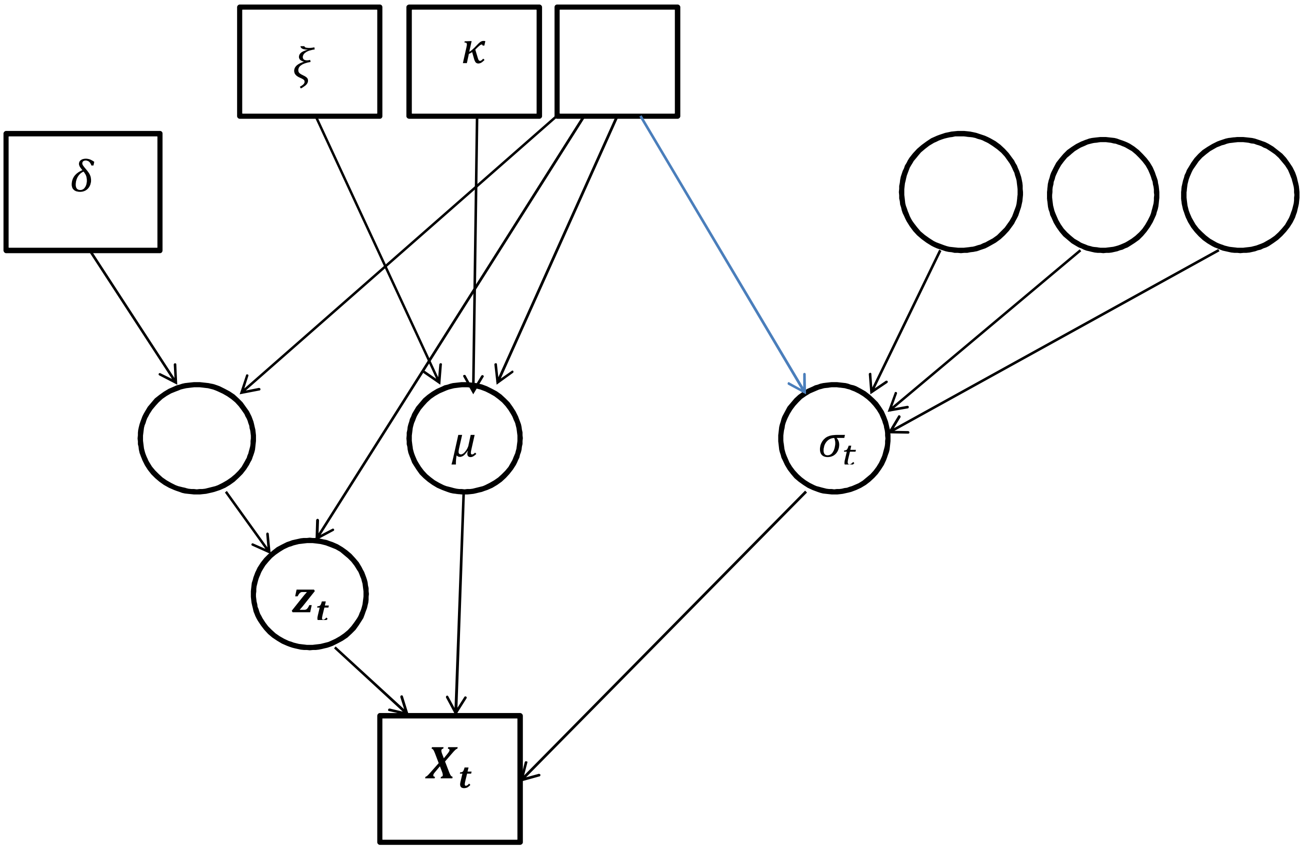


Fig. 2. Directed Acyclic Graph to the complete Hierarchical Model

For the distribution of β_t based on hierarchical NM (K) - GARCH (1, 1) models, we shall use five move types.

- Updating the weights p= $(p_1, p_2, \ldots\ldots, p_k)$
  Through conjugacy, the full conditional for the weights p remains Dirichlet in form:

$$p^{(m)} | \ldots\ldots \sim D(\delta + n_1^{(m-1)}, \ldots\ldots, \delta + n_K^{(m-1)}) \quad (21)$$

Where $n_j^{(m)} = \sum_{t=1}^{T} Z_{ij}^{(m)}$ is the number of observations currently allocated to the j component of the normal mixture. Here and the rest of the paper, '|' denotes 'conditional on all other variables'.

- Updating the parameters $\mu = (\mu_1, \mu_2, \ldots\ldots \mu_K)$
  The full conditions for μ_j are

$$\mu_j^{(m)} | \ldots\ldots \sim N\left(\frac{\sum_{t:z_t^{(m-1)}=j}\left(\frac{x_t}{\sigma_{j,t}^{-2(m-1)}}\right)+\kappa\xi}{\sum_{t:z_t^{(m-1)}=j}\left(\frac{1}{\sigma_{j,t}^{-2(m-1)}}\right)+\kappa}, (\sum_{t:z_t^{(m-1)}=j}\left(\frac{1}{\sigma_{j,t}^{-2(m-1)}}\right)+\kappa)^{-1}\right)$$

(22)

In order to preserve the ordering constraint on the μ_j, the move is accepted provided the ordering is unchanged and rejected otherwise.

- Updating the parameters (w_j, a_j, b_j), j=1…K.

The posterior conditional density for (w_j, a_j, b_j) are

$$p\left(w_j^{(m)}, a_j^{(m)} b_j^{(m)} \middle| \ldots\right) \propto \prod_{t:z_{ij}^{(m-1)}=1} \emptyset_j(y_t | Z_{t-1}) \times I(w_j > 0, a_j \geq 0, b_j \geq 0, a_j + b_j < 1)$$

(23)

We update (w_j, a_j, b_j) independently by using the Metropolis–Hastings (MH) algorithm.

- Updating the allocating z_{tj}

For the allocations we have the conditional probability

$$\mathrm{f}\,(z_{tj}^{(m)} = 1 | \ldots.) \propto \frac{p_j^{(m)}}{\sigma_j^{(m)}} \exp\{-\frac{\left(x_i - \mu_j^{(m)}\right)^2}{2\sigma_j^{2(m)}}\}$$

(24)

We can sample directly from this distribution and update the allocation variables independently through Gibbs sampling.

- Updating RRA β_t

β_t must satisfy the following equation

$$\mathrm{g}(\beta_t^{(m)} | \Im_{t-1}, \mu_j^{(m)}, \sigma_{j,t}^{2(m)}, j=1,\ldots,K) = \sum_{j=1}^{K} p_j^{(m)}\, e^{\mu_j^{(m)}\left(1-\beta_t^{(m)}\right)+\frac{\sigma_{j,t}^{2}\left(1-\beta_t^{(m)}\right)^2}{2}} - \sum_{j=1}^{k} p_j^{(m)}\, e^{-\mu_j^{(m)}\beta_t^{(m)}+\frac{\sigma_{j,t}^{2(m)}\beta_t^{2(m)}}{2}} = 0$$

(25)

We update β_t independently by means of the Gibbs sampler.

3.2 Model 2 Mixture of Normal

In Model 1 when we give a=b=0, the model reduce to Mixture of two normal distributions. In this model, RRA is not changed by the time period and it is estimated by

$$\sum_{j=1}^{k} p_j\, e^{\mu_j(1-\beta)+\frac{\sigma_j^2(1-\beta)^2}{2}} - \sum_{j=1}^{k} p_j\, e^{-\mu_j\beta+\frac{\sigma_j^2\,\beta^2}{2}} = 0$$

Here we are considering k=2. RRA is estimated using MLE and Bayesian inference.

3.2.1 Maximum Likelihood Estimation

EM algorithm is used to for maximum likelihood estimation. Once the parameters are estimated RRA is estimated using above equation. R package MixTools is used for estimation.

3.2.2 Bayesian Estimation

Prior for the mean and proportion is same as in Model 1. For variance we assume prior as inverse gamma distribution.

$$\sigma_j^2 \sim \mathrm{IG}(a, b)$$

where a and b selected such that the prior is non informative prior.

3.3 Model 3 GARCH Model

In Model 1 when we consider only one population, i.e. k=1, p_1=1, this reduces to Simple GARCH model.

Here we are considering GARCH (1, 1). RRA is estimated using MLE and Bayesian inference.

$$\widehat{\beta}_t = \frac{u}{\sigma_t^2} + 0.5$$

3.3.1 Maximum likelihood estimation

Once the parameters are estimated RRA is estimated using above equation. R package fGarch is used for estimation

3.3.2 Bayesian Estimation

Prior for the mean and the variance process is same as in Model 1

4 Simulation Studies

Basing on all the three Models, we perform simulation studies to compare the proposed MLE and Bayesian RRA estimations.

4.1 Model 1 NM-GARCH

We have included the simulation of the model and Bayesian program. See Appendix Model1 R Codes, where GARCH_SIM generates the simulated data. BUGS program for Bayesian estimation has been included in Appendix Model 1 OpenBugs code. We have skipped the estimation part.

4.2 Model 2 Mixture of Normal

In the Model2, we generate data from the Mixture model with five different sample sizes, N=50, 100, 200, 500, 1000. The true parameters are (p_1, p_2)= (0.3, 0.7), (μ_1, μ_2)= (−0.014, 0.014), and(σ_1^2, σ_2^2) = (0.003, 0.001). Thus, these values imply the true RRA parameter, β_{true}=3.564. Table 1 displays the average and standard deviation of RRA estimators as well as their mean square errors, $\boldsymbol{MSE}(\widehat{\boldsymbol{\beta}}) = [\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}})]^2 + \boldsymbol{Var}(\widehat{\boldsymbol{\beta}})$, across 200 samples. We see that MLE is good estimates in terms of bias, but Bayesian estimator is good in terms of MSE. However, as the sample size increases the bias is substantially reduced for all estimation methods. Based on the MSE criterion, the RRA estimated by Bayesian approach more accurate in all samples.

MLE estimated using R, MixTools package is used. The R function "NormalMixEM" is used for the estimation.

Table 1. Simulation results from MLE and Bayesian

N	β_{true}	MLE			Bayesian		
		$\widehat{\beta}$	$SE(\widehat{\beta})$	$MSE(\widehat{\beta})$	$\widehat{\beta}$	$SE(\widehat{\beta})$	$MSE(\widehat{\beta})$
50	3.5644	4.44136	3.87041	15.7491	1.96453	1.20767	**4.01805**
100	3.5644	3.67673	2.46669	6.09717	2.53666	1.00043	**2.05712**
200	3.5644	3.6454	1.73865	3.02947	2.94379	1.0663	**1.52216**
500	3.5644	3.60179	1.11818	1.25172	3.36222	0.91495	**0.87801**
1000	3.5644	3.53551	0.81618	0.66698	3.57013	0.73083	**0.53414**

The table considers N=50, 100, 200, 500, and 1000 observations obtained by generating sample from NM (2) models with true parameter values (p_1, p_2)= (0.3, 0.7), (μ_1, μ_2)= (−0.014, 0.014), and(σ_1^2, σ_2^2) = (0.003, 0.001). Therefore, from Eq. (15), we have the true RRA, β True=3.564 By simulating 200 times, the means, $\hat{\beta}$, standard deviations,$SE(\hat{\beta})$, and mean square errors, $MSE(\hat{\beta})$, of RRA estimators for two different estimating methods are shown. Bayesian Estimation has been done with number of iteration =20,000, burning period=6000 and thin =2. The bold figure represents the minimum MSE among MLE and Bayesian,

4.3 Model 3 GARCH Model

In the model 3, we generate data from the mixture model with five different sample sizes N= 25, 50,100. The true parameter are $\mu = 0.014 \times 0.3$ and(w, a, b) = (8.45E-05, 0.1, 0.85). Thus, these values imply the true RRA parameter, β_{True}=2.984. Table 2 displays the average and standard deviation of RRA estimators as well as their mean square errors,

$$\boldsymbol{MSE}(\hat{\beta}) = [\mathrm{Bias}(\hat{\beta})]^2 + \mathrm{Var}(\hat{\beta})$$

across 200 samples. We see that MLE is good estimates in terms of bias, but Bayesian estimator is good in terms of MSE. However, as the sample size increases the bias is substantially reduced for all estimation methods. Based on the MSE criterion, the RRA estimated by Bayesian approach can be seen to be more accurate in all samples. MLE has been estimated using R, fGarch package. The R function “GarchFit” has been used for the estimation.

Table 2 Simulation results from MLE and Bayesian

N	β_{true}	MLE			Bayesian		
		$\hat{\beta}$	$SE(\hat{\beta})$	$MSE(\hat{\beta})$	$\hat{\beta}$	$SE(\hat{\beta})$	$MSE(\hat{\beta})$
25	2.984	3.151437	6.728746	45.30406	2.662745	4.194065	17.69339
50	2.984	3.290328	3.968289	15.84116	2.269986	2.588269	7.208951
100	2.984	3.639503	3.486881	12.58802	2.621263	2.289852	5.375

The table considers N= 25, 50,100 observations obtained by generating sample from GARCH models with true parameter values $\mu = 0.014 \times 0.3$ and(w, a, b) = (8.45E-05, 0.1, 0.85). Therefore, from Eq. (15), we have the true RRA, β True=2.984. By simulating 200times, the means $\widehat{\beta}$, standard deviations $SE(\hat{\beta})$, and mean square errors$MSE(\hat{\beta})$, of RRA estimators for two different estimating methods are shown. Bayesian Estimation has been done with number of iteration =10,000, burning period=3000 and thin =2.

5 Empirical Studies

5.1 Dynamic RRA Estimation

The empirical analyses in this paper are based on monthly market rates of return and riskless rates during the period January 1941 through December 2001 with a sample of T=732 Observations. The value-weighted index of the New York Stock Exchange from 1941 to 2001 are collected from the Bloomberg database and used as the proxies for the rates of return on the market portfolio. The proxies for the risk-free rates are three monthly returns of U.S. Treasury bills (a one month treasury bill is considered in the paper). Using this set of data, we calculate the excess return = (1 + Rmt)/ (1 + Rft), the return in the market portfolio over the risk-free rate, for doing the following empirical studies. Table „Descriptive Statistics" provides the summary statistics for the whole period of the log excess returns which are based on the monthly value-weighted index and the 30-day U.S. Treasury bills returns from January 1941 to December 2010. The summary statistics includes sample mean, minimum, median, maximum, standard deviation, skewness, kurtosis, the p-values of chi-square and Jarque and Bera (1980)2 tests to check the normality of the data, and the p-values of Ljung and Box (1978) to verify the serial correlations of returns and square of returns data.

From the Table, we can observe that the unconditional distribution of the log excess return has negative skewness and heavy tails relative to the normal distribution. For chi-square and Jarque and Bera (1980) tests, we reject the normality assumption for overall periods at 1% significant level. Therefore, it appears that normal distribution is not an adequate assumption to the log transformation for the excess returns. In addition, the results of Ljung-Box test confirm that log monthly excess returns of the value-weighted index for overall period have no significant serial correlations. The table also shows that there are some serial correlations for square of the log returns and there is volatility clustering phenomenon. Therefore, assuming log monthly excess returns to be GARCH (1, 1) model may be more appropriate. Therefore, the MLE and Bayesian estimators of dynamic RRA based on the GARCH (1, 1) model from 1/ 1970 to 12/2001 are shown in Figs. 3.

5.2 Descriptive statistics

Descriptive statistics for monthly log excess returns (%) of value-weighted index (1/1941–12/2010)

A. Summary statistics for log excess returns (%)			
Mean	0.2724	S.D.	4.158871
Minimum	-25.07	Skewness	-0.5882245
Median	0.5333	Kurtosis	5.104726
Maximum	14.43		
B. Test for Normality			
Jarque–Bera test		Chi-square test	
J–B stat.	179.2978	Chi-square stat.	39.377
J–B p-value	$< 2.2e^{-16}$ ***	Chi-square p-value	0.03376 **
C. Ljung–Box test for log excess returns, y_t			
Q(5) stat.	7.9906	Q(5) p-value	0.1568
Q(10) stat.	9.1528	Q(10) p-value	0.5177
D. Ljung –Box test for square of log excess returns, y_t^2			
Q(5) stat.	13.5388	Q(5) p-value	0.01882 **
Q(10) stat.	24.1321	Q(10) p-value	0.007257 ***

The table shows the summary statistics of log (monthly excess returns), y_t over the period from 1/1941 to 12/2010 in panel A. The statistics and p-value of chi-square and Jarque and Bera (1980) tests are used to check the closeness of the data to a normal density in panel B. Moreover, we exhibit the p-value of Ljung-Box Q-statistic (1978) to measure the serial correlations of y_t and y_t^2 using five and ten lagged values. ** and *** indicate statistical significance at the 5% and 1% levels (two tailed test), respectively.

5.3 Diagnostics Tests

The diagnostic plot for the Bayesian estimates are provided in the. The comments on the graphs are given below.

Trace plots & History plots

The trace plot for the parameter “a” is showing evidence of stationarity. But “w”,”b”,”mu” charts are showing non stationary behavior.

Auto Correlation plot

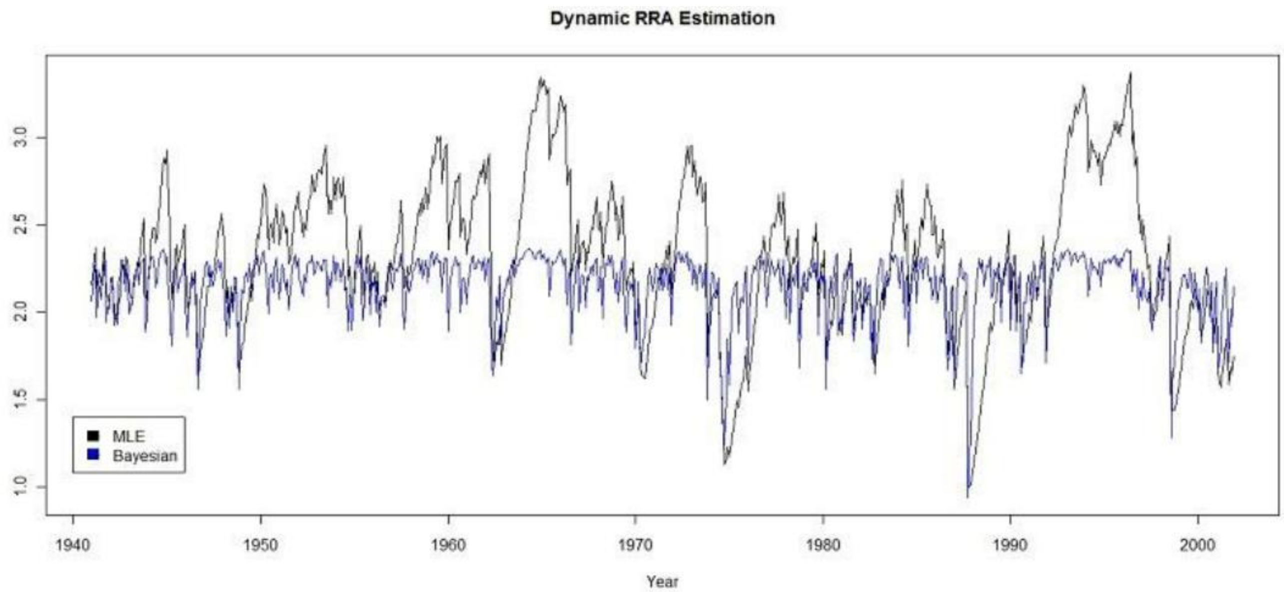
Similar inference can also be seen in an auto correlation plot. Increase the thin value can reduce the autocorrelation.

5.4 GARCH model estimation

	Estimate	Std. error	t value	Pr (>|t|)
mu	2.953e-03	1.48E-03	1.999	0.04566
omega	1.123e-04	5.35E-05	2.097	0.03599
alpha1	6.585e-02	2.21E-02	2.976	0.00292
beta1	8.711e-01	3.87E-02	22.53	<2e-16

	Mean	SD	MC_error	val2.5pc	Median	val97.5pc
a (alpha1)	0.093	4.83E-02	0.000529	2.05E-02	0.0853	0.2092
B (beta1)	0.419	1.95E-01	0.008811	0.04387	0.4235	0.8373
mu	0.002	2.39E-03	0.0000455	-0.00303	0.001855	0.006528
W (omega)	0.001	4.04E-04	0.0000181	1.90E-04	0.001026	0.001838

5.5 Dynamic RRA Estimation



In the above figure, the RRA estimated by MLE method, on average, is larger than that estimated by Bayesian method. In addition, we find that Bayesian estimation method is less sensitive than MLE estimation method.

6 Conclusion/ Future Research

We had spent significant amount of time to replicate NM-GARCH (1, 1) by trying different prior distribution/reformulation. But the technical difficulty led us to go with simpler models.

The analysis in the research paper by WU & Lee (2007) is based on 1 month T-Bill. We have used 3 months T-Bill which brought differences in the summary of the paper and our work. This paper shows that the utility-based model of asset pricing can be estimated more accurately with the GARCH (1, 1) model for log asset returns.

The MLE and Bayesian approach with a weakly informative prior have been derived to estimate RRA. As per our empirical findings, the log excess returns can be adequately characterized by the GARCH (1, 1) model with their excess kurtosis.

Also it has been identified that RRA estimator is statistically efficient with the robust model assumption. The RRA estimator obtained by the Bayesian approach with a weakly informative prior performs better in small samples based on the MSE criterion. Finally, Bayesian approach can combine an investor's prior belief about the accuracy of the pricing model and the information in the data and describe the sampling distribution of RRA estimator.

7 Appendix

7.1 R CODES

7.1.1 Model 1 R codes

Function used for generate simulated Mixed GARCH data

```
GARCH_SIM <-function (m,p,g1,g2,n,seed)
{
# g = garch part
# n = number of observation
# seed = random seed
set.seed( seed)
b <- rbinom(n,1,p)
mn <- p*m[1]+ (1-p)*m[2]
e <- rnorm(n)
x <- double(n)
h <- double(n)
if(b[1]==1)
{ h[1] <- g1[1]/(1.0-g1[2]-g1[3]) }
else
{ h[1] <- g2[1]/(1.0-g2[2]-g2[3]) } # Need to replace by Overall Variance.
x[1] <- rnorm(1,m[2-b[1]], sd = sqrt( h[1]))
for(i in 2:n) # Generate GARCH(1,1) process
{
if(b[i]==1) {
h[i]=g1[1]+g1[2]*(x[i-1]-mn)^2 +g1[3]*h[i-1]
x[i] <- m[1] + e[i]*sqrt(h[i]) }
else {
h[i]=g2[1]+g2[2]*(x[i-1]-mn)^2 +g2[3]*h[i-1]
x[i] <- m[2] + e[i]*sqrt(h[i]) }
}
print(c(mean(x),mn))
print(c(var(x),p*g1[1]+(1-p)*g2[1]+ p*(1-p)*(m[1]-m[2])^2 ))
return(x)
}
#y=GARCH_SIM(c(100,-100),.3,c( 1,0,0),10000,1)
#plot(density(y))
m=c(-0.014,0.014)
g1=c(0.00015,0.1,0.85)
g2=c(0.00005,0.1,0.85)
y=GARCH_SIM(m ,.3,g1,g2 ,1000,1)
```

7.1.2 Model 2 R Codes

Simulation of Normal Mixture

```
# Simulation Code for Mixture of Two Normal.
normix <- function(n, M, S, p) {
u <- runif(n)
distr = ifelse(u <= p, 1, 2)
mu <- M[distr]
sigma<- sqrt(S [distr])
rnorm (n, mean=mu, sd=sigma)
}
RRA function.
betaf=function(b,prop,mu,sigma)
{
#p =.3
#m= -0.014
#s1= 0.00015/(1-.95)
#s2= 0.00005/(1-.95)
p=prop[1]
m1=mu[1]; m2=mu[2];
s1=sigma[1]^2; s2=sigma[2]^2;
f1= exp( m1*(1-b)+s1*(1-b)^2/2 ) - exp( -m1*b + s1*b^2/2 )
f2= exp( m2*(1-b)+s2*(1-b)^2/2 ) - exp( -m2*b + s2*b^2/2 )
f3 = p*f1+(1-p)*f2
return(f3)
}
```

MLE estimation of RRA

```
library(mixtools)
y=normix(n,c(-0.014,0.014),c(0.003,0.001),.3)
mle=normalmixEM(y,k=2)
mlep=mle$lambda
mlem= mle$mu
mles=mle$sigma
uniroot(betaf,c(-20,20), prop=mlep,mu=mlem,sigma=mles)$root
```

7.1.3 Model 3 R Codes

Simulation of GARCH

```
library(fGarch)
spec = garchSpec(model = list(mu = 0.001, omega = 1e-6, alpha = 0.1, beta = 0.8))
y=garchSim(spec, n )
MLE estimation of RRA
mle= garchFit(~ garch(1,1), data = y, trace = FALSE)@fit$matcoef[,1]
```

7.1.4 Empirical Analysis

```
y=log(x)
yt <- ts(y, start=c(1952, 10), frequency=12)
#ts.plot(yt)
library(moments)
summary(y)
sd(y)
skewness(y)
kurtosis(y)
library(fGarch)
pchiTest(y)
jarqueberaTest(y)
Box.test (y, lag = 5, type="Ljung")
Box.test (y, lag = 10, type="Ljung")
Box.test (y^2, lag = 5, type="Ljung")
Box.test (y^2, lag = 10, type="Ljung")
mlegg= garchFit(~ garch(1,1), data = y, trace = FALSE)
mle=mlegg@fit$matcoef[,1]
mlemu=mle[1]
mleg= mle[2:4]
betaf=function(beta,mu,g)
{
#mu= -0.014
#g= c(1e-6, 0.1, 0.8)
m1=mu;
#w=g[1]; a=g[2];b=g[3];
s=g[1]/(1-g[2]-g[3])
f1= exp( m1*(1-beta)+s*(1-beta)^2/2 ) - exp( -m1*beta + s*beta^2/2 )
return(f1)
}
betaf1=function(beta,m1,s)
{
f1= exp( m1*(1-beta)+s*(1-beta)^2/2 ) - exp( -m1*beta + s*beta^2/2 )
return(f1)
}
btf=function(m,s)
{
return(uniroot(betaf1,c(-200,200),m1=m,s=s)$root)
}
ht=Betat=0;
ht[1]=mleg[1]/(1-mleg[2]-mleg[3])
Betat[1]=btf(mlemu,ht[1])
for(i in 2:length(mlegg@residuals))
{
ht[i]= mleg[1] + mleg[2]*mlegg@residuals[i-1]^2 +mleg[3]*ht[i-1]
Betat[i]=btf(mlemu,ht[i])
}
```

7.2 OPEN BUGS CODES

MODEL 1

```
model{
y[1]~ dnorm(mu[1],tau[1] )
mu[1] <- lambda[T[1]]
tau[1]<- 1/ht[1,T[1]]
ht[1,1] <- w1/(1-a1-b1)
ht[1,2]<- w2/(1-a2-b2)
T[1] ~ dcat(P[])
for( i in 2:N)
{
y[i]~ dnorm(mu[i],tau[i] )
mu[i] <- lambda[T[i]]
tau[i]<- 1/ht[i,T[i]]
ht[i,1] <- w1+a1*(y[i-1]-mu[i-1] )*(y[i-1]-mu[i-1] )+b1*ht[i-1,T[i-1]]
ht[i, 2]<- w2+a2*(y[i-1]-mu[i-1] )*(y[i-1]-mu[i-1] )+b2*ht[i-1,T[i-1]]
T[i] ~ dcat(P[])
} P[1:2] ~ ddirich(alpha[])
theta ~ dunif(0.0, 1000)
lambda[2] <- lambda[1] + theta
lambda[1] ~ dnorm(0.0, 1.0E-6)
w1 ~ dunif(0.0,1)
a1~dunif(0.0,1)
b11~ dunif(0.0,1)
b1<-(1-a1)*b11
w2 ~ dunif(0.0,1)
a2~dunif(0.0,1)
b22~ dunif(0.0,1)
b2<-(1-a2)*b22
```

MODLE 2

```
model
{
for( i in 1 : N ) {
y[i] ~ dnorm(m[i], t[i])
m[i] <- mu[T[i]]
t[i] <-tau[T[i]]
T[i] ~ dcat(P[])
}
P[1:2] ~ ddirch(alpha[])
theta ~ dnorm(0.0, 10)I(0.0, )
mu[2] <- mu[1] + theta
mu[1] ~ dnorm(0.0, 1)
tau[2] ~ dgamma(40, 0.1) sigma[2] <- 1 / sqrt(tau[2])
tau[1] ~ dgamma(40, 0.1) sigma[1] <- 1 / sqrt(tau[1])
}
```

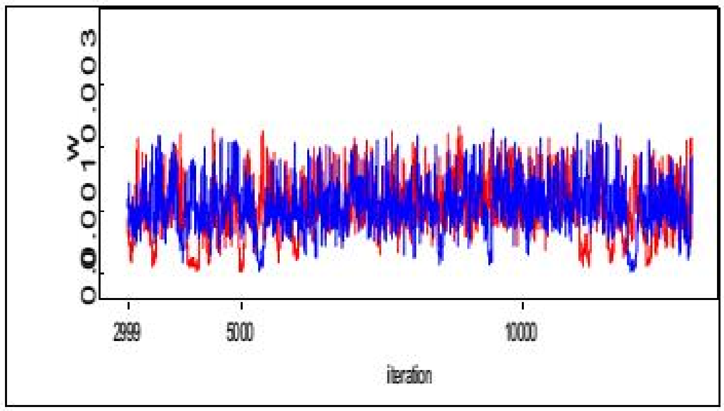
25

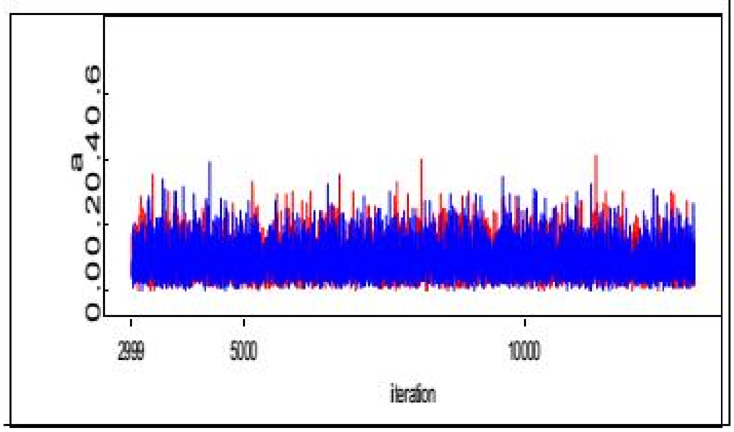
MODEL 3

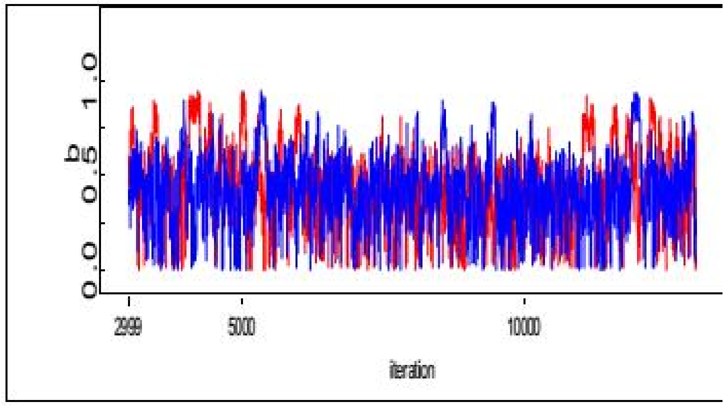
```
model
{
y[1] ~ dnorm(mu, t[1])
t[1]<- 1 / s2[1]
s2[1]<- w / (1-a-b)
e[1]<-y[1]-mu
beta[1] <-mu/s2[1]+ 0.5
for( i in 2 : N ) {
y[i] ~ dnorm(mu, t[i])
t[i]<- 1 / s2[i]
s2[i]<-w+ a*(y[i-1]-mu)*(y[i-1]-mu)+ b*s2[i-1]
e[i]<-y[i]-mu
beta[i] <-mu/s2[i]+ 0.5
}
mu ~ dnorm(0, 1)
w ~ dunif(0,1)
a1~dunif(0.0,1)
b1~ dunif(0.0,1)
c1~ dunif(0.0,1)
a<-a1/(a1+b1+c1)
b<-b1/(a1+b1+c1)
}
```

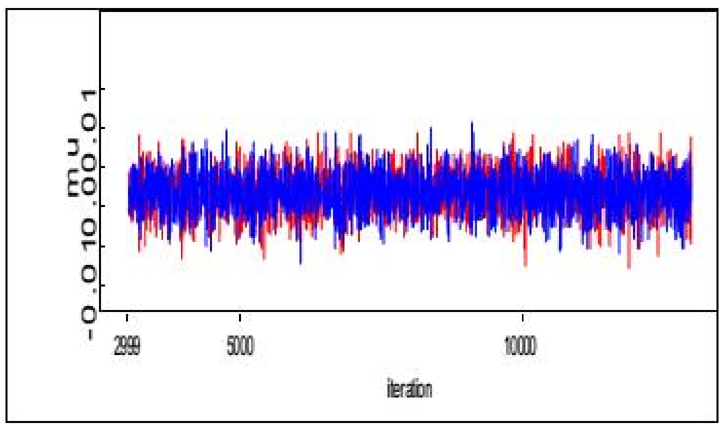
7.3 Diagnostics Tests

7.3.1 History Plot

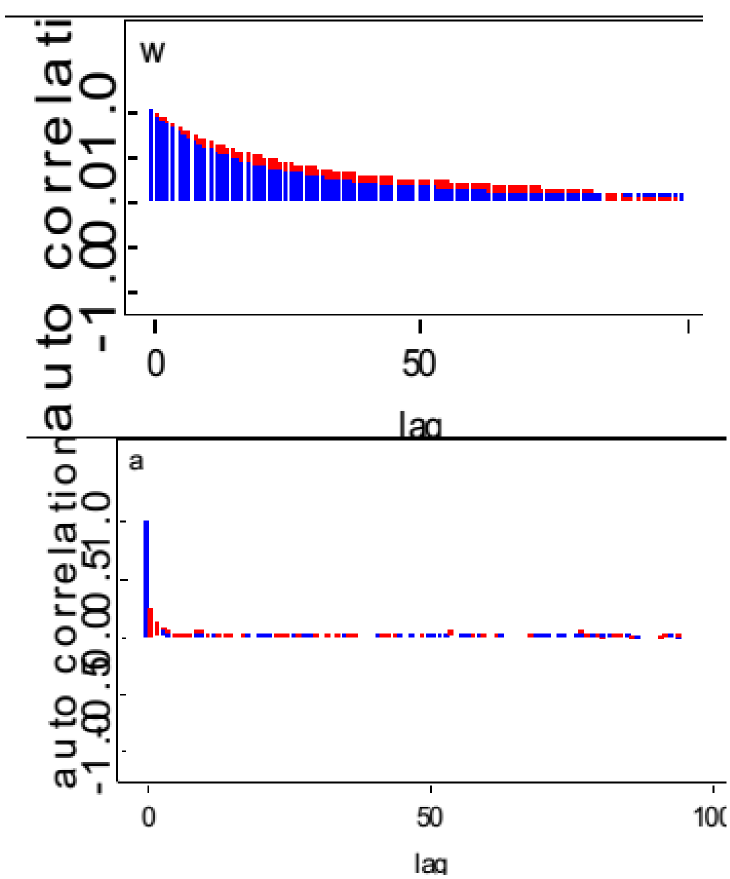


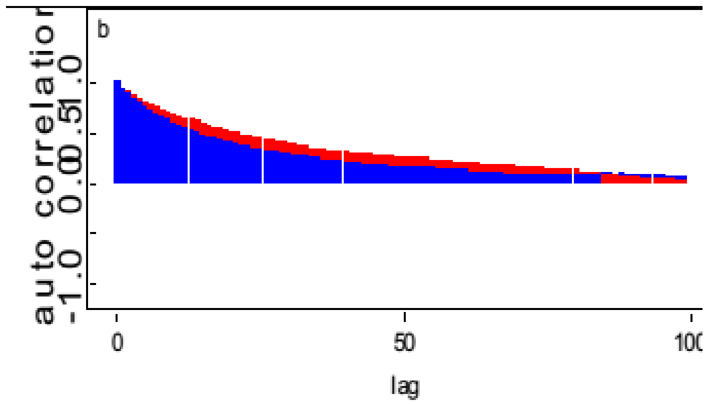


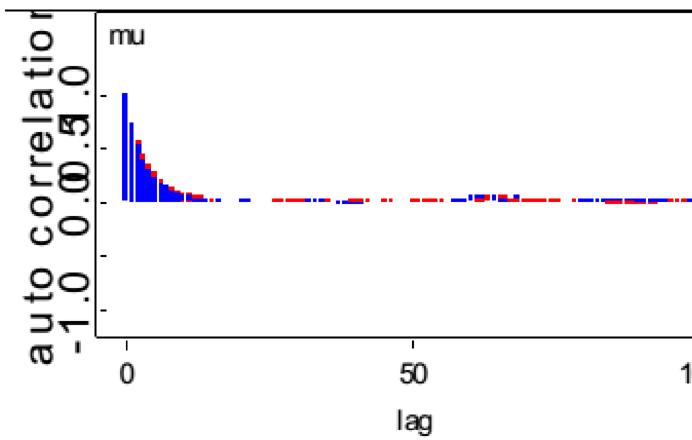

b
0.0
0.5
1.0
2999
5000
10000
iteration


mu
-0.010
-0.010
0.00
0.01
2999
5000
10000
iteration

7.3.2 Auto-correlation Plot

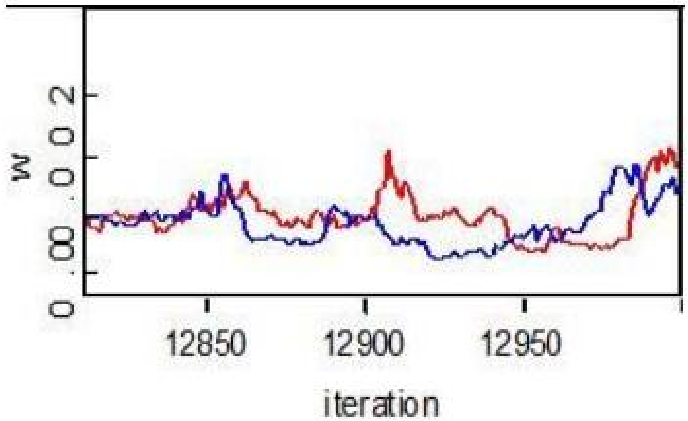


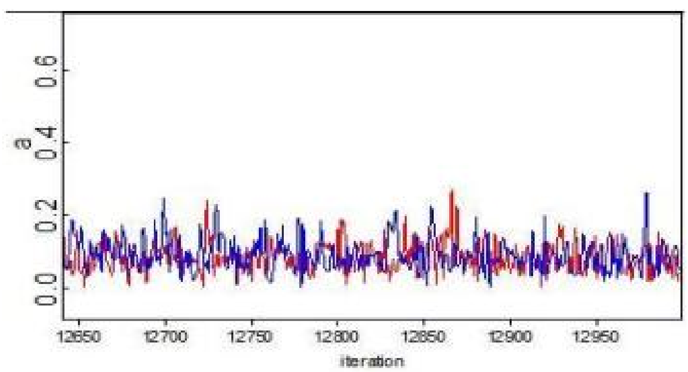

b
auto correlation
-1.0
0.0
0.5
1.0
0
50
100
lag

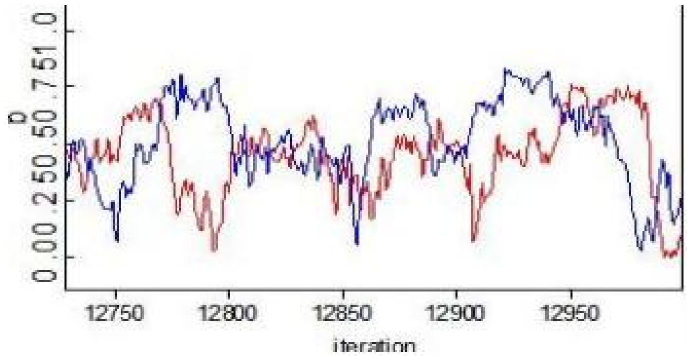

mu
auto correlation
-1.0
0.0
0.5
1.0
0
50
lag

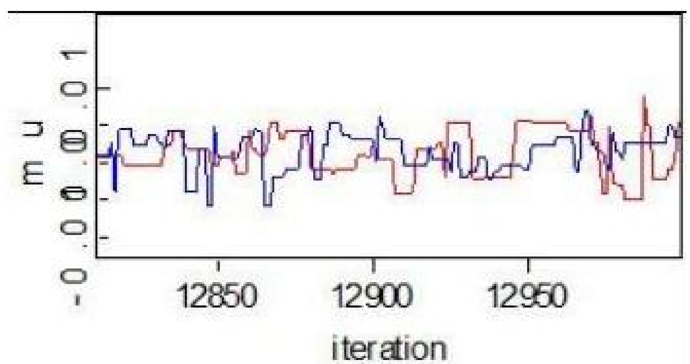
29

7.3.3 Trace Plot









8 References

Wu, C., & Lee, J. C. (2007). Estimation of a utility-based asset pricing model using normal mixture GARCH (1, 1). *Economic Modelling, 24*(2), 329-349.