STATISTICAL ANALYSIS OF FINANCIAL TIME SERIES AND RISK MANAGEMENT

by
Hongyu Ru

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operatiaons Research (Statistics).

Chapel Hill
2012

Approved by:

Eric Ghysels

Chuanshu Ji

Amarjit Budhiraja

Shankar Bhamidi

Riccardo Colacito

Abstract

HONGYU RU: STATISTICAL ANALYSIS OF FINANCIAL TIME SERIES AND RISK MANAGEMENT.
(Under the direction of Eric Ghysels.)

The dissertation studies the dynamic of volatility, skewness, and value at risk for financial returns. It contains three topics.

The first one is the asymptotic properties of the conditional skewness model for asset pricing. We start with a simple consumption-based asset pricing model, and make a connection between the asset pricing model and the regularity conditions for a quantile regression. We prove that the quantile regression estimators are asymptotically consistent and normally distributed under certain assumptions for the asset pricing model.

The second one is about dynamic quantile models for risk management. We propose a financial risk model based on dynamic quantile regressions, which allows us to estimate conditional volatility and skewness jointly. We compare this approach with ARCH-type models by simulation. We also propose a density fitting approach by matching conditional quantiles and parametric densities to obtain the conditional distributions of returns.

The third one is a simulation study of a consumption based asset pricing model. We show that larger returns and Sharp ratio can be obtained by introducing conditional asymmetry in the asset pricing model.

Acknowledgments

Being a graduate student, there is nothing else more exciting than writing the acknowledgement in my dissertation, as it is one of the best opportunities for me to extend my heartfelt gratitude and deepest appreciation to all who make this possible over the years.

First of all, I would like to thank my dissertation advisor Professor Eric Ghysels for his extraordinary support and encouragement. His unparalleled enthusiasm, dedication and vision on research always give me inspiration to keep going on my work.

I would also like to sincerely thank Professor Chuanshu Ji who also supervised my research. It has been a hard journey for me to pursue this degree. Whenever I was in trouble, he has always tried his best to help me out on both research and life throughout the years.

Moreover, I am very grateful to other member of my dissertation committee: Professor Amarjit Budhiraja, Professor Riccardo Colacito, Professor Shankar Bhamidi and Professor Eric Renault for their fruitful discussions and stimulations that led me to finish this dissertation.

Finally, my heartfelt thanks goes to my family, especially my parents, sister and husband, for their tremendous support through the ups and downs of my life.

Table of Contents

List of Figures . . . . **vii**

List of Tables . . . . **viii**

1 Asymptotic Properties of Quantile-based Conditional Skewness Models for Asset Pricing **1**

1.1 Introduction 1

1.2 The Asset Pricing Model 4

1.3 The Empirical Quantile Model 6

1.3.1 A robust measure of conditional asymmetry 7

1.3.2 Conditional quantile specification and estimation 9

1.4 Asymptotic Properties 10

1.5 Conclusion 14

1.6 Proofs 14

2 Dynamic Quantile Models for Risk Management **24**

2.1 Introduction 24

2.2 The Generic Setup 26

2.3 Dynamic Quantile Models 29

2.4 Quantile Distribution Fits 33

2.5 Simulation 35

2.5.1 Simulation of Conditional Heteroskedasticity versus Quantils . . 35
2.6 Conclusion 42
2.7 Tables and Figures 42

3 Simulation Study of Long Run Skewness for Asset Pricing 53
3.1 Introduction 53
3.2 Model Specification and Calibration 54
3.2.1 Model Specification 54
3.2.2 Calibration 57
3.3 Simulation 57
3.3.1 Hansen and Jagannathan Bound 57
3.3.2 Equity Returns 58
3.3.3 Conditional Moments 59
3.4 Conclusion 60
3.5 Tables and Figures 61

Bibliography 70

List of Figures

2.1 HYBRID quantile regression and MIDAS quantile regression 50

2.2 Comparison of quantiles by quantile distribution fits and CAViaR model 51

2.3 Comparison of Expected Shortfall(ES) by quantile distribution fits and regression based ES of CAViaR quantiles 52

3.1 Conditional Moments of x_t for multiple horizons: moments of $x_{t+1}|x_t$ in blue, $x_{t+3}|x_t$ in green and $x_{t+12}|x_t$ in red 68

3.2 Conditional Moments of excess return for multiple horizons: moments of $r_{e,t+1}|x_t$ in blue, $r_{e,t+3}|x_t$ in green and $r_{e,t+12}|x_t$ in red 69

List of Tables

2.1 Hybrid quantiles and MIDAS quantiles for 5% VaR 43

2.2 Hybrid quantiles and MIDAS quantiles for 1% VaR 44

2.3 Summary of Model Specifications 45

2.4 Summary of Parameters in Simulation Study 46

2.5 Comparison of σ_t using QLIKE 47

2.6 Comparison of σ_t using MSEprop 48

2.7 Comparison of VaR using MSE 49

3.1 Monthly Calibration 62

3.2 Distribution of Predictive Components for Monthly Calibration 62

3.3 Equity return for $\gamma = 15$ 63

3.4 Equity return for $\gamma = 10$ 64

3.5 Multihorizon equity return for $\gamma = 15$ 65

3.6 Multihorizon equity return for $\gamma = 15$ 66

3.7 Stochastic discount factor 67

Chapter 1

Asymptotic Properties of Quantile-based Conditional Skewness Models for Asset Pricing

1.1 Introduction

It has been documented by empirical studies that the distribution of stock market returns, either conditional or unconditional, can not be fully characterized by just mean and variance. Many previous studies have shown that the stock market returns are negatively skewed(see e.g. Harvey and Siddique (2000)). Researchers begin to incorporate the third moment - skewness, into financial models and applications. One of the applications of using skewness is portfolio selection. Harvey and Siddique (2000) has discussed about investors' preference on the skewness of a portfolio. A portfolio with positive skewness is preferred by investors if everything else is equal. But all those results are subjected to the robustness of the measure of skewness due to the following reasons.

Stock market returns, especially in emerging markets, are known to have fat tails. The conventional measures of the moments are based on sample averages. Therefore, those estimators are sensitive to outliers, especially for the third and higher moments. To study the stock market returns more accurately, researchers in financial areas begin to seek for robust measures that are less sensitive to outliers (see e.g. Kim and White

(2004)). Kim and White (2004) has surveyed several more robust measures of skewness based on quantiles and moments, which have been originally introduced by statisticians(see, e.g. Bowley (1920)). But those are only unconditional skewness measures. To study the dynamics of the stock market returns or financial time series, we need a robust measure for conditional skewness.

White, Kim, and Manganelli (2008) have proposed a conditional version for the measure introduced by Bowley (1920) by replacing the unconditional quantiles with conditional quantiles. To estimate conditional quantiles, we need back to the definition of regression quantile. Regression quantile has been first introduced by Koenker and Bassett (1978), which extended sample quantiles to linear regression quantiles. They defined a minimization problem, and defined the solution to that minimization problem as regression quantile. White (1996) has made an important contribution by proving the consistency of the nonlinear regression quantiles for stationary dependent cases. Another important contribution to the estimation of conditional quantiles was made by Weiss (1991). In this paper, the author has introduced a least absolute error estimator, which is a special case of regression quantiles, for dynamic nonlinear models with non i.i.d. errors. The author shows that the estimator is consistent and asymptotically normal under some regularity conditions and has also provided an estimator for asymptotic covariance matrix. Engle and Manganelli (2004) have applied nonlinear regression quantiles to study the dynamic of value at risk, which is a quantile. The authors have proved that the estimator is consistent and asymptotically normal under some regularity conditions, and provided an estimator for asymptotic covariance matrix for nonlinear conditional quantiles in the context of time series. White, Kim, and Manganelli (2008) have extended this method and estimated multiple quantiles jointly.

The quantile regression models used in White, Kim, and Manganelli (2008) are for one-period return. Ghysels, Plazzi, and Valkanov (2010a) have proposed a quantile regression model that can be used for n-period, long-horizon return based on daily information. They find that conditional skewness still varies across time even for GARCH- and TARCH-filtered returns. In this chapter, we focus on the quantile regression models of Ghysels, Plazzi, and Valkanov (2010a).

The asypototic properties of those conditional quantile models have been studied by several papers(see, e.g., White, Kim, and Manganelli (2008), Engle and Manganelli (2004)). They show that the conditional quantile estimators are consistent and asymptotically normal under some regulation conditions. But those regulation conditions are hard to be verified empirically. Motivated by the limitation of those regularity conditions, we are seeking from modeling the data generating process(DGP) from an asset pricing model to derive the regularity conditions of the quantile regression model of Ghysels, Plazzi, and Valkanov (2010a). In other words, we want to construct the link between those regulation conditions proposed by White, Kim, and Manganelli (2008), and Engle and Manganelli (2004) and basic DGPs with some simple assumptions.

Now, the question is what DGP is a good model for the economy and can generate a fairly decent amount of time-varying conditional skewness like what we have observed in the real data (Ghysels, Plazzi, and Valkanov (2010a)). Campbell and Cochrane (1999) have presented a consumption-based asset pricing model that can explain important asset market phenomena. In addition, the model can produce non-normal consumption-based stock prices and returns with negative skewness. Bansal and Yaron (2004) have also presented a consumption-based asset pricing model which includes a long-run predictable component. Their model can also explain some key features of dynamic asset pricing phenomena. But for these two models, they don't have analytical solutions for the price-dividend ratio and returns, which are needed for constructing the connection between DGP and regularity conditions for quantile regression. Burnside (1998) has provided an asset pricing model with normal shocks to consumption growth.

Tsionas (2003) has extended Burnside (1998) to allow for any shock that has moment generating functions. Both of them have analytical solution for price-dividend ratio, and therefore returns. Tsionas (2003) can generate conditional skewness,[1] but we don't know if it can create time-varying conditional skewness. Bekaert and Engstrom (2010) may be another option, which has both analytical solutions and allows for time-varying conditional skewness for consumption growth.[2]

In this paper, we start with a rather conventional asset pricing framework based on discounted dividend streams. Initially we use closed-form formulas of Burnside (1998) and Tsionas (2003) using first a Gaussian setting and subsequently a general setting that allows us to characterize DGP's for which we subsequently study the asymptotic properties of conditional quantile regressions and skewness measures. We have proved that the conditional quantile estimators are consistent and asymptotically normalunder those simple assumptions for the DGP of asset pricing we use.

This chapter is structured as follows. Section 1.2 describes the asset pricing model. Section 1.3 describes the quantile regression model. In Section 1.4, we explore the asympototic properties of quantile regression under the assumed data generating process. Section 1.5 concludes this chapter and describes the future works. Regulation conditions and proofs are in Section 1.6.

1.2 The Asset Pricing Model

First order condition of asset pricing to price an asset that entitles a dividend D_t in each period satisfy

$$P_t = E_t \left[S_{t,t+1}(P_{t+1} + D_{t+1}) \right],$$

[1]For example, if the shock distribution is a general Edgeworth expansion, then it allows for skewness.

[2]But we don't know if we can prove all the regularity conditions under this model, since they assume the parameter for shocks follow AR(1) process, namely the shocks are dependent.

where P_t is price of the asset at time t, $S_{t,t+1}$ is stochastic discount factor(SDF). We consider a representative agent with CRRA preference and denote the price-dividend ratio as $v_t = P_t/D_t$, then we have

$$v_t = E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+v_{t+1})\frac{D_{t+1}}{D_t}\right], \tag{1.1}$$

where γ is the coefficient of relative risk aversion, β is the discount factor, and C_t is the consumption at time t. Assume the log dividend growth $x_t = \log(C_{t+1}/C_t) = \log(D_{t+1}/D_t)$ follows AR(1) process

$$x_t = (1-\rho)\mu + \rho x_{t-1} + \xi_t, \tag{1.2}$$

where ρ is the persistent parameter, and ξ_t is an i.i.d sequence of random variables.

Assumption 1 *(i)* $|\rho| < 1$ *and* $\rho \neq 0$*;*

(ii) Let $M_{\xi_t}(s) \equiv E\exp(s\xi_t)$ *be the moment generating function(MGF) of* ξ_t*,* $M_{\xi_t}(s)$ *exists;*

(iii) Let $f_{\xi_t}(\xi_t)$ *be the probability density of* ξ_t*,* $f_{\xi_t}(\xi_t)$ *is everywhere continuous, continuously differentiable and* $f_{\xi_t}(\xi_t) > 0$.

The unconditional distribution of x_t is $\mu + (1-\rho)^{-1}\xi_t$ and MGF of x_t is $M_{x_t}(s) = \exp(\mu s)M_{\xi_t}(s/(1-\rho))$. Tsionas (2003) shows that

$$v_t = \sum_{i=1}^{\infty}\beta^i \exp\left[a_i + b_i(x_t-\mu)\right] \equiv \sum_{i=1}^{\infty} z_i, \tag{1.3}$$

where $\alpha \equiv 1-\gamma$, $\theta \equiv (1-\gamma)/(1-\rho)$

$$a_i = \alpha i\mu + \sum_{j=1}^{i}\log M_{\xi_t}(\theta(1-\rho^j))$$

$$b_i = \alpha \frac{\rho}{1-\rho}(1-\rho^i).$$

The conditions for stationary and bounded equilibrium to exist are given by Tsionas (2003).

Assumption 2 *Let* $r \equiv \beta \exp(\alpha\mu) M_{\xi_t}(\theta)$, $r < 1$.

Lemma 1 *Under Assumption 1, 2,*

(i) the series v_t *converges;*

(ii) the series v_t *have finite moments of every integer order.*

Proof: See Tsionas (2003).

We are now in position to study the property of the returns generated from this asset pricing model. The log return can be expressed as

$$r_{t+1} = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) = \log(1 + v_{t+1}) - \log v_t + x_{t+1}. \tag{1.4}$$

Lemma 2 $E\left|r_t\right|^3 < \infty$ *if Assumption 1, and Lemma 1 holds.*

Proof: See Section 1.6.

Given Assumption 1 and 2, it is possible to show that the series of returns have finite moments of every integer order. Here we just show that the series of returns have finite third moments, which is sufficient for our latter use. The proofs for the returns to have higher order moments are similar.

1.3 The Empirical Quantile Model

The setup of the empirical quantile models follows Ghysels, Plazzi, and Valkanov (2010a) closely. In section 1.3.1, we describe the robust measure of conditional asymmetry. In Section 1.3, we present the conditional quantile regression specification and the estimation of the model.

A robust measure of conditional asymmetry

In section 1.2, the returns generated from the DGP's are one-period return, which can be daily, weekly, or monthly, etc. We are interested in the asymmetry in the conditional distributions of n-period returns. Let $r_{t,n} = \sum_{j=0}^{n-1} r_{t+j}$, for $n \geq 2$, be the log continuously compounded n-period return of an asset, where r_t is the one-period log return. Let $F_n(r) = P\left(r_{t,n} < r\right)$ be the unconditional cumulative distribution function (CDF) of $r_{t,n}$, and $F_{n,t|t-1}\left(r\right) = P\left(r_{t,n} < r|I_{t-1}\right)$ be the conditional CDF given the information set I_{t-1}. The θth quantile can be defined as

$$q^{*}_{\theta_k}\left(r_{t,t+n}\right) \equiv \inf\left\{r : F_n\left(r\right) = \theta_k\right\}, \ \theta_k \in (0,1].$$

If $F_n(r)$ and $F_{n,t|t-1}\left(r\right)$ are strictly increasing, then the θth quantile of return $r_{t,n}$ is

$$q_{\theta}\left(r_{t,n}\right) = F_n^{-1}\left(r\right), \ \theta \in (0,1]$$

and the conditional θth quantile of return $r_{n,t}$ is

$$q_{\theta,t}\left(r_{n,t}\right) = F^{-1}_{n,t|t-1}\left(r\right), \ \theta \in (0,1]. \tag{1.5}$$

For the sake of simplicity, we could assume that $F_n(r)$ and $F_{n,t|t-1}\left(r\right)$ are strictly increasing such that the inverse of $F_n\left(r\right)$ or $F_{n,t|t-1}\left(r\right)$ is unique. Later in the next section, we are going to show that strictly increasing can be verified under standard regularity conditions.

As discussed in Section 1.1, researches have proposed robust measures of asymmetry other than sample average to estimate skewness. Bowley (1920) is one of them.

Bowley's (1920) robust coefficient of skewness is defined as

$$CA\left(r_{t,n}\right) = \frac{\left(q_{0.75}\left(r_{t,n}\right) - q_{0.50}\left(r_{t,n}\right)\right) - \left(q_{0.50}\left(r_{t,n}\right) - q_{0.25}\left(r_{t,n}\right)\right)}{q_{0.75}\left(r_{t,n}\right) - q_{0.25}\left(r_{t,n}\right)} \tag{1.6}$$

where $q_{0.25}\left(r_{t,n}\right)$, $q_{0.50}\left(r_{t,n}\right)$ and $q_{0.75}\left(r_{t,n}\right)$ are the 25th, 50th, and 75th unconditional quantiles of $r_{t,n}$.

Groeneveld and Meeden (1984) have proposed four properties that any reasonable skewness measure should satisfy. That is for skewness measure $\gamma\left(y_t\right)$ (See Kim and White (2004)):

(i) for any $a > 0$ and b, $\gamma\left(y_t\right) = \gamma\left(ay_t + b\right)$;

(ii) if y_t is symmetric, then $\gamma\left(y_t\right) = 0$;

(iii) $-\gamma\left(y_t\right) = \gamma\left(-y_t\right)$;

(iv) if F and G are cumulative distribution function of y_t and x_t, and $F <_c G$, then $\gamma\left(y_t\right) \leq \gamma\left(x_t\right)$, where $<_c$ is a skewness-ordering among distribtutions.

The measure (1.6) satisfies all the four conditions (See Groeneveld and Meeden (1984)). Also this measure is normalized to be unit independent with values between -1 and 1. The negative(positive) values of this measure indicate skewness to the left(right). Although this measure is robust, it is an unconditional skewness measure, which can not be used to study the dynamics of conditional asymmetry and those properties of financial time series.

Recently, White, Kim, and Manganelli (2008) and Ghysels, Plazzi, and Valkanov (2010a) have used a conditional version of (1.6) given information I_{t-1}, which makes studying the dynamics of conditional asymmetry using a measure like (1.6) possible.

They define

$$CA_t\left(r_{t,n}\right) = \frac{\left(q_{0.75,t}\left(r_{t,n}\right) - q_{0.50,t}\left(r_{t,n}\right)\right) - \left(q_{0.50,t}\left(r_{t,n}\right) - q_{0.25,t}\left(r_{t,n}\right)\right)}{q_{0.75,t}\left(r_{t,n}\right) - q_{0.25,t}\left(r_{t,n}\right)}. \tag{1.7}$$

where $q_{0.25,t}\left(r_{t,n}\right)$, $q_{0.50,t}\left(r_{t,n}\right)$ and $q_{0.75,t}\left(r_{t,n}\right)$ are the 25th, 50th, and 75th conditional quantiles of $r_{t,n}$. To estimate (1.7), we need estimate the conditional quantiles of $r_{t,n}$. In the next section, we present our models and estimation methods for those conditional quantiles in (1.7).

Conditional quantile specification and estimation

We denote the θth conditional quantile of $r_{t,n}$ at time t as $q_{\theta,t}\left(r_{t,n};\delta_{\theta,n}\right)$, where $\delta_{\theta,n}$ is the vector of parameters to be estimated for θth quantile at horizon n. Denote the information set that contains the daily information up to time $t-1$ as $I_{t-1} = \{x_{t-1}, x_{t-2}, ...\}$, where x_t is a vector of daily conditioning variables. We use a mixed data sampling (MIDAS) approach to setup the model for conditional quantile of $r_{t,n}$, which are multiple horizon returns, based on daily returns in the information set I_{t-1}. In other words, we use daily returns as regressors. The model is defined as follows

$$q_{\theta,t}\left(r_{t,n};\delta_{\theta,n}\right) = \alpha_{\theta,n} + \beta_{\theta,n} Z_t\left(\kappa_{\theta,n}\right) \tag{1.8}$$

$$Z_t\left(\kappa_{\theta,n}\right) = \sum_{d=1}^{D} w_d\left(\kappa_{\theta,n}\right) x_{t-d} \tag{1.9}$$

where $\delta_{\theta,n} = \left(\alpha_{\theta,n}, \beta_{\theta,n}, \kappa_{\theta,n}\right)'$ are unknown parameters to estimate. Following Ghysels, Santa-Clara, and Valkanov (2006), we specify $\omega_d\left(\kappa_{\theta,n}\right)$ as

$$\omega_d\left(\kappa_{\theta,n}\right) = \frac{f\left(\frac{d-1/2}{D}, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)}{\sum_{m=1}^{D} f\left(\frac{m-1/2}{D}, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)}, \tag{1.10}$$

where $\kappa_{\theta,n} = (\kappa_{1,\theta,n}, \kappa_{2,\theta,n})$ is a 2-dimensional row vector that reduces the number of weights for lag coefficient to estimate from D to 2, $f(z,a,b) = z^{a-1}(1-z)^{b-1}/\beta(a,b)$, $\beta(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, and Γ is Gamma function. We specify the daily return x_{t-d} in (2.15) as $|r_{t-d}|$.

We estimate the parameters $\delta_{\theta,n}$ in (2.14-1.10) with non-linear least squares. More specifically, for a given quantile θ and horizon n, we minimize

$$\min_{\delta_{\theta,n}} T^{-1} \sum_{t=1}^{T} \rho_{\theta,n}(\varepsilon_{\theta,n,t}) \tag{1.11}$$

where $\varepsilon_{\theta,n,t} = r_{t,n} - q_{t,n}(\theta; \delta_{\theta,n})$, $\rho_{\theta,n}(\varepsilon_{\theta,n,t}) = (\theta - 1\{\varepsilon_{\theta,n,t} < 0\})\varepsilon_{\theta,n,t}$ is the usual "check" function used in quantile regressions. If the model we specified is the true model of DGP, and $\delta_{\theta,n}$ are true unknown parameters, then $Q_{\theta,n}(\varepsilon_{\theta,t}|I_{t-1}) = 0$, where $Q_{\theta,n}(\varepsilon_{\theta,t}|.)$ is the θ conditional quantile of $\varepsilon_{\theta,n,t}$. The soluction to the optimization problem (1.11) can also be considered as quasi-maximum likelihood estimator (QMLE), where $\rho_{\theta,n}(\varepsilon_{\theta,n,t})$ is the log-likelihood of independent asymmetric double exponential random variable which belongs to tick-exponential family (see e.g. White, Kim, and Manganelli (2008), and Komunjer (2004)).

1.4 Asymptotic Properties

The asymptotic properties of $\hat{\delta}_{\theta,n}$ that minimizes (1.11) have been studied by several papers(see e.g. White (1996), Weiss (1991), Engle and Manganelli (2004) and White, Kim, and Manganelli (2008)). They have shown that the estimates $\hat{\delta}_{\theta,n}$ are consistent and asymptotically normal by assuming that the DGP satisfied some regularity conditions. But those regulation conditions are hard to be verified empirically. Motivated by the limitation of those regularity conditions, we are seeking from modeling the data generating process(DGP) from a basic asset pricing model to derive the regularity conditions of the quantile regression model of Ghysels, Plazzi, and Valkanov (2010a).

We consider the data are generated by DGP described in Section 1.2 and estimate the conditional quantiles using models described in Section 1.3. First, we define some properties for the parameter space. Then, we prove all the assumptions (see White, Kim, and Manganelli (2008)) that are needed for consistency and asymptoticly normality under our DGP of asset pricing models described in Section 1.2. To fix notation, all the following statements are for fixed n and fixed θ.

Assumption 3 *Let the parameter space* $\tilde{A} \equiv \{\delta_{\theta,n} : \beta_{\theta,n} \neq 0, \kappa_{1,\theta,n} > 0, \kappa_{2,\theta,n} > 0\}$ *be a compact subset of* R^4*, and* A *be a compact subset of* $\tilde{A}$*. Assume that the true parameter* $\delta^0_{\theta,n} \in A$ *and* $\delta^0_{\theta,n} \in int\,(A)$*.*

Lemma 3 *Let* Ω *be the sample space. Under Assumption 3, the function* $q_{\theta,t}(\omega, \delta_{\theta,n})$ *is such that*

(i) for each t *and each* $\omega \in \Omega$*,* $q_{\theta,t}(\omega, \cdot)$ *is continuous, continuously differentiable, twice continuously differentiable on* A*;*

(ii) for each t *and each* $\delta_{\theta,n} \in A$*,* $q_{\theta,t}(\cdot, \delta_{\theta,n})$*,* $\nabla q_{\theta,t}(\cdot, \delta_{\theta,n})$*, and* $\nabla^2 q_{\theta,t}(\cdot, \delta_{\theta,n})$ *are* I_{t-1} *measurable, where* $\nabla q_{\theta,n}(\cdot, \delta_{\theta,n})$ *denote the gradient(row vector) of scaler function* $q_{\theta,n}(\cdot, \delta_{\theta,n})$ *with respect to* $\delta_{\theta,n}$*.*

Proof: See Section 1.6.

Lemma 4 *For fixed* θ *and* $\delta_{\theta,n}$*,* $E|r_{t,t+n}|$*,* $E|q_{\theta,t}|$*, and* $E|\varepsilon_{\theta,t}|$ *are finite on* A *if Assumption 3 and Lemma 2 hold.*

Proof: See Section 1.6.

Lemma 5 *Let* $D_{0,t} \equiv \sup_{\delta_{\theta,n} \in A} |q_{\theta,t}(\cdot, \alpha_{\theta,n})|$*,* $D_{1,t} \equiv \max_{i=1,\ldots,4} \sup_{\delta_{\theta,n} \in A} \left|\partial_{\delta_{i,\theta,n}} q_{\theta,t}(\cdot, \delta_{\theta,n})\right|$*, and* $D_{2,t} \equiv \max_{i=1,\ldots,4} \max_{j=1,\ldots,4} \sup_{\delta_{\theta,n} \in A} \left|(\partial_{\delta_{i,\theta,n}} \partial_{\delta_{j,\theta,n}} q_{\theta,t}(\cdot, \delta_{\theta,n})\right|$*, where* $\delta_{i,\theta,n}$ *is the ith* *component of* $\delta_{\theta,n}$*. Under Assumption 3, if Lemma 2 holds, then (i)* $E\left(D_{0,t}\right)<\infty$*; (ii)* $E(D_{1,t}^{3})<\infty$ *; (iii)* $E(D_{2,t}^{2})<\infty$*.*

Proof: See Section 1.6.

Lemma 6 $\{\rho_{\theta,n}(\varepsilon_{\theta,t})\}$ *is strictly stationary and ergodic, and obeys the uniform law of large number, if Lemma 4 and Lemma 5(i) hold.*

Proof: See Section 1.6.

Lemma 7 *Let* $h_{\theta,t}\left(r_{t,n}|I_{t-1}\right)$ *be the conditional density of* $r_{t,n}$ *given* I_{t-1}*. Under Assumption 1,*

(i) for each θ *and each* t*,* $h_{\theta,t}\left(r_{t,n}|I_{t-1}\right)$ *is everywhere continuous;*

(ii) for each θ *and each* t*,* $h_{\theta,t}\left(r_{t,n}|I_{t-1}\right)>0$*;*

(iii) there exists a finite positive constant N *such that for each* θ*, and each* t*,* $h_{\theta,t}\left(r_{t,n}|I_{t-1}\right)\leq N<\infty$*;*

(iv) there exists a finite positive constant L *such that for each* θ*, each* t*, and each* $\lambda_1,\lambda_2\in\mathbb{R}$*,* $\left|h_{\theta,t}\left(\lambda_1|I_{t-1}\right)-h_{\theta,t}\left(\lambda_2|I_{t-1}\right)\right|\leq L\left|\lambda_1-\lambda_2\right|$*.*

Proof: See Section 1.6.

Lemma 8 *For fixed* t *and every* $\tau>0$*, there exists* $\delta_\tau>0$ *such that for all* $\delta_{\theta,n}\in A$ *with* $\left\|\delta_{\theta,n}-\delta_{\theta,n}^{0}\right\|>\tau$*,* $P\left(\left|q_{\theta,t}(\cdot,\delta_{\theta,n})-q_{\theta,t}(\cdot,\delta_{\theta,n}^{0})\right|>\delta_\tau\right)>0$ *if Lemma 10 holds.*

Proof: See Section 1.6.

Lemma 9 *Let* $Q^0\equiv E\left[h_{\theta,t}\left(0|I_{t-1}\right)\nabla q_{\theta,t}'\left(\cdot,\delta_{\theta,n}^0\right)\nabla q_{\theta,t}\left(\cdot,\delta_{\theta,n}^0\right)\right]$ *and* $V^0\equiv E\left(\eta_{\theta,t}^{0\prime}\eta_{\theta,t}^{0}\right)$*, where* $\eta_{\theta,t}^0\equiv\nabla q_{\theta,t}\left(\cdot,\delta_{\theta,n}^0\right)\psi_\theta\left(\varepsilon_{\theta,t}\right)$ *and* $\psi_\theta\left(\varepsilon_{\theta,t}\right)\equiv\theta-1_{\left\{\varepsilon_{\theta,t}<0\right\}}$*. If Lemma 10 and 7 hold, then (i)* Q^0 *is positive definite; (ii)* V^0 *is positive definite.*

Proof: See Section 1.6.

Now, we are in position to have the results of consistency and asymptoticly normality.

Theorem 1 *If Assumption 3, Lemma 3, 4, 5(i), 6 - 8 hold, then* $\hat{\delta}_{\theta,n} \xrightarrow{a.s} \delta^0_{\theta,n}$.

Proof: See White, Kim, and Manganelli (2008).

Theorem 2 *If Assumption 3, Lemma 3 - 9 hold, then*

$$\sqrt{T} V^{0-1/2} Q^0 \left(\hat{\delta}_{\theta,n} - \delta^0_{\theta,n} \right) \xrightarrow{d} N(0, I).$$

Proof: See White, Kim, and Manganelli (2008).

The consistent estimators for V^0 and Q^0 have been given by several papers(see e.g. White, Kim, and Manganelli (2008) and Engle and Manganelli (2004)) with one additional assumption.

Theorem 3 *Let* $\hat{V}_T \equiv T^{-1} \sum_{t=1}^{T} \hat{\eta}'_t \hat{\eta}_t$, $\hat{\eta}_t \equiv \nabla q_{\theta,t} \left(\cdot, \hat{\delta}_{\theta,n} \right) \psi_\theta \left(\hat{\varepsilon}_{\theta,t} \right)$, $\hat{\varepsilon}_{\theta,t} \equiv r_{t,t+n} - q_{\theta,t} \left(\cdot, \hat{\delta}_{\theta,n} \right)$. *If Assumption 3, Lemma 3 - 9 hold, then* $\hat{V}_T \xrightarrow{p} V^0$.

Proof: See White, Kim, and Manganelli (2008).

Assumption 4 $\{\hat{c}_T\}$ *is a stochastic sequence and* c_T *is a nonstochastic sequence such that (i)* $\hat{c}_T / c_T \xrightarrow{p} 1$*; (ii)* $c_T = o(1)$*; (iii)* $c_T^{-1} = o\left(T^{1/2}\right)$.

Theorem 4 *Let* $\hat{Q}_T = (2\hat{c}_T T)^{-1} \sum_{t=1}^{T} 1_{-\hat{c}_T \le \hat{\varepsilon}_{\theta,t} \le \hat{c}_T} \nabla' q_{\theta,t} \left(\cdot, \delta_{\theta,n} \right) \nabla q_{\theta,t} \left(\cdot, \delta_{\theta,n} \right)$. *If Assumption 3, 4, Lemma 3 - 9 hold, then* $\hat{Q}_T \xrightarrow{p} Q^0$.

Proof: See White, Kim, and Manganelli (2008).

1.5 Conclusion

In this chapter, we start with a simple consumption-based asset pricing model with CRRA utility, and make a connection between the asset pricing model and the regularity conditions for a quantile regression, which is hard to be verified. We prove that the quantile regression estimators are asymptotically consistent and normally distributed under certain assumptions for the asset pricing model.

1.6 Proofs

This section contains the proofs for this chapter.

Proof of Lemma 2: We show $Er_{t+1}^2 < \infty$ by showing that $E\left|r_{t+1}\right|^3 < \infty$. Since $v_{t+1} > 0$, we have $0 < \log(1 + v_{t+1}) < v_{t+1}$,

$$\begin{aligned}
E\left|r_{t+1}\right|^3 \leq & E\left|\log\left(1+v_{t+1}\right)\right|^3 + E\left|\log v_t\right|^3 + E\left|x_{t+1}\right|^3 + 3E\left|\log\left(1+v_{t+1}\right)\left(\log v_t\right)^2\right| \\
& + 3E\left|\left(\log\left(1+v_{t+1}\right)\right)^2 \log v_t\right| + 3E\left|\left(\log\left(1+v_{t+1}\right)\right)^2 x_{t+1}\right| \\
& + 3E\left|\left(\log\left(1+v_{t+1}\right)\right) x_{t+1}^2\right| + 3E\left|\left(\log v_t\right)^2 x_{t+1}\right| \\
& + 3E\left|\left(\log v_t\right) x_{t+1}^2\right| + 6E\left|\left(\log\left(1+v_{t+1}\right)\right)\left(\log v_t\right) x_{t+1}\right| \\
\leq & Ev_{t+1}^3 + E\left|\log v_t\right|^3 + E\left|x_{t+1}\right|^3 + 3E\left|v_{t+1}\left(\log v_t\right)^2\right| + 3E\left|v_{t+1}^2 \log v_t\right| \\
& + 3E\left|v_{t+1}^2 x_{t+1}\right| + 3E\left|v_{t+1} x_{t+1}^2\right| + 3E\left|\left(\log v_t\right)^2 x_{t+1}\right| \\
& + 3E\left|\left(\log v_t\right) x_{t+1}^2\right| + 6E\left|v_{t+1}\left(\log v_t\right) x_{t+1}\right| \\
\leq & E\left|v_{t+1}\right|^3 + E\left|\log v_t\right|^3 + E\left|x_{t+1}\right|^3 + 3\left(E\left|v_{t+1}\right|^3\right)^{\frac{1}{3}}\left(E\left|\log v_t\right|^3\right)^{\frac{2}{3}} \\
& + 3\left(E\left|v_{t+1}\right|^3\right)^{\frac{2}{3}}\left(E\left|\log v_t\right|^3\right)^{\frac{1}{3}} + 3\left(E\left|v_{t+1}\right|^3\right)^{\frac{2}{3}}\left(E\left|x_{t+1}\right|^3\right)^{\frac{1}{3}} \\
& + 3\left(E\left|v_{t+1}\right|^3\right)^{\frac{1}{3}}\left(E\left|x_{t+1}\right|^3\right)^{\frac{2}{3}} + 3\left(E\left|\log v_t\right|^3\right)^{\frac{2}{3}}\left(E\left|x_{t+1}\right|^3\right)^{\frac{1}{3}} \\
& + 3\left(E\left|\log v_t\right|^3\right)^{\frac{1}{3}}\left(E\left|x_{t+1}\right|^3\right)^{\frac{2}{3}} + 6\left(E\left|v_{t+1}\right|^3 E\left|\log v_t\right|^3 E\left|x_{t+1}\right|^3\right)^{\frac{1}{3}}
\end{aligned}$$

The last inequlity holds due to Holder's inequality. We know that $E\left|v_{t+1}\right|^3 < \infty$ and $E\left|x_{t+1}\right|^3 < \infty$ from Lemma 1. Now we need to show $E\left|\log v_t\right|^3 < \infty$ to have $E\left|r_{t+1}\right|^3 < \infty$. Considering the negative part of $\left(\log v_t\right)^3$, since $z_i > 0$, $\log z_i \leq \log \sum_{i=1}^{\infty} z_i$, we have

$$\left[\left(\log v_t\right)^3\right]^- = \left[\left(\log \sum_{i=1}^{\infty} z_i\right)^3\right]^- \leq \left[\left(\log z_1\right)^3\right]^-,$$

where $\log z_1 = \log\beta + a_1 + b_1(x_t - \mu) = \log\beta + a_1 + b_1\left(1-\rho\right)^{-1}\xi_t$. Since the unconditional distribution of x_t is given by $x_t = \mu + \left(1-\rho\right)^{-1}\xi_t$(see Tsionas (2003)). By the assumption that the MGF of ξ exists, all the moments of ξ exists. Hence, $E\left(\log z_1\right)^3 < \infty$, $E\left|\log z_1\right|^3 < \infty$ and $E\left(\left(\log z_1\right)^3\right)^- < \infty$. $\left(-\log v_t\right)^3$ is convex because

$(-\log v_t)$ is convex and $g(x) = x^3$ is convex and nondecreasing. Hence, $(\log v_t)^3$ is concave. Thus, $E(\log v_t)^3 \leq (\log Ev_t)^3 < \infty$. Therefore,

$$E\left[(\log v_t)^3\right]^+ = E(\log v_t)^3 + E\left[(\log v_t)^3\right]^- \leq (\log Ev_t)^3 + E\left[(\log z_1)^3\right]^- < \infty$$

$$E|logv_t|^3 = E\left[(\log v_t)^3\right]^+ + E\left[(\log v_t)^3\right]^- < \infty$$

It follows that $E|r_{t+1}|^3 < \infty$. ■

Proof of Lemma 3: Let $z_d \equiv \frac{d-1/2}{D}$, and $g(z,a,b) \equiv z^{a-1}(1-z)^{b-1}$, we have

$$\omega_d(\kappa_{\theta,n}) = \frac{g(z_d, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}$$

$$\partial_{\kappa_{1,\theta,n}}\omega_d(\kappa_{\theta,n}) = (\kappa_{1,\theta,n}-1)\,\omega_d(\kappa_{\theta,n})\left[z_d^{-1} - \frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}-1, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}\right]$$

$$\partial_{\kappa_{2,\theta,n}}\omega_d(\kappa_{\theta,n}) = (\kappa_{2,\theta,n}-1)\,\omega_d(\kappa_{\theta,n})\left[(1-z_d)^{-1} - \frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}-1)}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}\right]$$

$$\begin{aligned}
\partial^2_{\kappa_{1,\theta,n}}\omega_d(\kappa_{\theta,n}) &= \omega_d(\kappa_{\theta,n})\left[z_d^{-1} - \frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}-1, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}\right] \\
&+ (\kappa_{1,\theta,n}-1)^2\left[z_d^{-1} - \frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}-1, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}\right]^2 \\
&+ (\kappa_{1,\theta,n}-1)^2\,\omega_d(\kappa_{\theta,n})\left[\frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}-1, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}\right]^2 \\
&- (\kappa_{1,\theta,n}-1)(\kappa_{1,\theta,n}-2)\,\omega_d(\kappa_{\theta,n})\frac{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}-2, \kappa_{2,\theta,n})}{\sum_{m=1}^{D} g(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n})}
\end{aligned}$$

$$
\begin{aligned}
&\partial^2_{\kappa_{1,\theta,n}} \omega_d \left(\kappa_{\theta,n}\right) = \omega_d \left(\kappa_{\theta,n}\right) \left[\left(1 - z_d\right)^{-1} - \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 1\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)} \right] \\
&+ \left(\kappa_{2,\theta,n} - 1\right)^2 \left[\left(1 - z_d\right)^{-1} - \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 1\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)} \right]^2 \\
&+ \left(\kappa_{2,\theta,n} - 1\right)^2 \omega_d \left(\kappa_{\theta,n}\right) \left[\frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 1\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)} \right]^2 \\
&- \left(\kappa_{2,\theta,n} - 1\right) \left(\kappa_{1,\theta,n} - 2\right) \omega_d \left(\kappa_{\theta,n}\right) \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 2\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)}
\end{aligned}
$$

$$
\begin{aligned}
&\partial_{\kappa_{1,\theta,n}} \partial_{\kappa_{2,\theta,n}} \omega_d \left(\kappa_{\theta,n}\right) = \\
&- \left(\kappa_{1,\theta,n} - 1\right) \left(\kappa_{2,\theta,n} - 1\right) \omega_d \left(\kappa_{\theta,n}\right) \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n} - 1, \kappa_{2,\theta,n} - 1\right)}{\left(\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)\right)^2} \\
&+ \left(\kappa_{1,\theta,n} - 1\right) \left(\kappa_{2,\theta,n} - 1\right) \omega_d \left(\kappa_{\theta,n}\right) \\
&\times \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n} - 1, \kappa_{2,\theta,n}\right) \sum_{l=1}^{D} g\left(z_l, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 1\right)}{\left(\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)\right)^2} \\
&\left(\kappa_{1,\theta,n} - 1\right) \left(\kappa_{2,\theta,n} - 1\right) \omega_d \left(\kappa_{\theta,n}\right) \left[z_d^{-1} - \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n} - 1, \kappa_{2,\theta,n}\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)} \right] \\
&\times \left[\left(1 - z_d\right)^{-1} - \frac{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n} - 1\right)}{\sum_{m=1}^{D} g\left(z_m, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)} \right].
\end{aligned}
$$

It is clear that Lemma 3 is satisfied under Assumption 3. ■

Proof of Lemma 4:

$$
E|r_{t,t+n}| = E\left|\sum_{j=0}^{n-1} r_{t+j}\right| \leq \sum_{j=0}^{n-1} E|r_{t+j}| < \infty
$$

Since the parameter space is compact set by Assumption 3, we have

$$E|q_{\theta,t}| = E\left|\alpha_{\theta,n} + \beta_{\theta,n}\sum_{d=1}^{D}\omega_d(\kappa_{\theta,n})|r_{t-d}|\right| \le |\alpha_{\theta,n}| + |\beta_{\theta,n}|\sum_{d=1}^{D}\omega_d(\kappa_{\theta,n})E|r_{t-d}| < \infty$$

$$E|\varepsilon_{\theta,t}| = E|r_{t,t+n} - q_{\theta,t}| \le E|r_{t,t+n}| + E|q_{\theta,t}| < \infty$$

■

Lemma 10 *For fixed t and $\delta_{\theta,n} \in A$, the components of $\nabla q_{\theta,t}(\cdot, \delta_{\theta,n})$ are linearly independent of each other almost surely under Assumption 3.*

Proof of Lemma 10: we check if there is nontrival $a \equiv (a_1, a_2, a_3, a_4)'$ such that for fixed t and $\delta_{\theta,n} \in A$, and every possible outcome of $|r_{t-d}|$, $\nabla q_{\theta,t}(r_{t,n}, \delta_{\theta,n})\, a = 0$. Since

$$\nabla q_{\theta,t}(r_{t,n}, \delta_{\theta,n}) =$$
$$\left(1, \sum_{d=1}^{D}\omega_d(\kappa_{\theta,n})|r_{t-d}|, \beta_{\theta,n}\sum_{d=1}^{D}\partial_{\kappa_{1,\theta,n}}\omega_d(\kappa_{\theta,n})|r_{t-d}|, \beta_{\theta,n}\sum_{d=1}^{D}\partial_{\kappa_{2,\theta,n}}\omega_d(\kappa_{\theta,n})|r_{t-d}|\right).$$

This yields

$$\begin{aligned} a_1 + \sum_{d=1}^{D}\omega_d(\kappa_{\theta,n})|r_{t-d}|\big(&a_2 + a_3\beta_{\theta,n}(\kappa_{1,\theta,n} - 1)\left(z_d^{-1} - c_1\right) \\ &+ a_4\beta_{\theta,n}(\kappa_{2,\theta,n} - 1)\left((1 - z_d)^{-1} - c_2\right)\big) = 0 \end{aligned}$$

where c_1 and c_2 are function of $\kappa_{1,\theta,n}$ and $\kappa_{1,\theta,n}$, but do not depend on d. Since $\omega_d(\kappa_{\theta,n}) > 0$, and 1 and $|r_{t-d}|$, $d = 1, \cdots, D$, are linearly independent almost surely, then $a_1 = 0$ and $a_2 + a_3\beta_{\theta,n}(\kappa_{1,\theta,n} - 1)\left(z_d^{-1} - c_1\right) + a_4\beta_{\theta,n}(\kappa_{2,\theta,n} - 1)\left((1 - z_d)^{-1} - c_2\right) = 0$, $d = 1, \ldots, D$. If $\beta_{\theta,n} \neq 0$, $\kappa_{1,\theta,n} \neq 1$, $\kappa_{2,\theta,n} \neq 1$ and $D > 3$, the linear system of equations have no nontrival solution a such that $\nabla q_{\theta,t}(r_{t,n}, \delta_{\theta,n})\, a = 0$ identically. Lemma

10 then follows. ■

Proof of Lemma 5: For any $\delta_{\theta,n} \in A$, $E\left|q_{\theta,t}\left(\cdot,\delta_{\theta,n}\right)\right| < \infty$. Lemma 5(i) then follows.

Proof of Lemma 3 indicates that $\partial_{\kappa_{1,\theta,n}}\omega_d\left(\kappa_{\theta,n}\right)$ is finite for all $\delta_{\theta,n} \in A$. If Lemma 2 holds, then for every $\delta_{\theta,n} \in A$, we have

$$\begin{aligned}
&E\left|\partial_{\kappa_{1,\theta,n}}q_{\theta,t}\left(r_{t,n},\delta_{\theta,n}\right)\right|^3 \\
&= E\beta_{\theta,n}^3\sum_{d=1}^{D}\sum_{l=1}^{D}\sum_{m=1}^{D}\partial_{\kappa_{1,\theta,n}}\omega_d\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_l\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_m\left(\kappa_{\theta,n}\right)\left|r_{t-d}\right|\left|r_{t-l}\right|\left|r_{t-m}\right| \\
&= \beta_{\theta,n}^3\sum_{d=1}^{D}\sum_{l=1}^{D}\sum_{m=1}^{D}\partial_{\kappa_{1,\theta,n}}\omega_d\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_l\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_m\left(\kappa_{\theta,n}\right)E\left|r_{t-d}\right|\left|r_{t-l}\right|\left|r_{t-m}\right| \\
&\leq \beta_{\theta,n}^3\sum_{d=1}^{D}\sum_{l=1}^{D}\sum_{m=1}^{D}\partial_{\kappa_{1,\theta,n}}\omega_d\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_l\left(\kappa_{\theta,n}\right)\partial_{\kappa_{1,\theta,n}}\omega_m\left(\kappa_{\theta,n}\right) \\
&\times\left(E\left|r_{t-d}\right|^3E\left|r_{t-l}\right|^3E\left|r_{t-m}\right|^3\right)^{1/3} \\
&< \infty.
\end{aligned}$$

Proof of $E\left|\partial_{\delta_{i,\theta,n}}q_{\theta,t}\left(r_{t,n},\delta_{\theta,n}\right)\right|^3 < \infty$ for other components is similar. Since for all $\delta_{\theta,n} \in A$ and $i = 1,\ldots,4$, $E\left|\partial_{\delta_{i,\theta,n}}q_{\theta,t}\left(r_{t,n},\delta_{\theta,n}\right)\right|^3$ is finite , we can conclude that $E(D_{1,t}^3) < \infty$.

The proof of Lemma 5(iii) is the same as that of Lemma 5(ii). ■

Proof of Lemma 6: Since $\{x_t\}$ is AR(1) process with i.i.d shocks and $|\rho| < 1$, it is strictly stationary and ergodic. When a process is strictly stationary, then a measurable function of this process is also strictly stationary. Similary property holds for ergodicity. Both $r_{t,t+n}$, and $q_{\theta,t}$ are measurable function of x_t, so $\rho_{\theta,n}$ is strictly stationary and ergodic. It has been shown by White, Kim, and Manganelli (2008) that $|\rho_{\theta,n}|$ is dominated by $2(|r_{t,t+n}| + |D_{0,t}|)$. Using Theorem A.2.2 on the appendix of White (1996), $\rho_{\theta,n}$ obeys the uniform law of large number. ■

Proof of Lemma 7: First, find the expression for $h_{\theta,t}(r_{t,n}|I_{t-1})$ as a function of $f_{\xi_t}(\xi_t)$.

$$\begin{aligned} v_t &= \sum_{i=1}^{\infty} \beta^i \exp\left[a_i + b_i\left(x_{t-1} - \mu\right)\right] \\ &= \sum_{i=1}^{\infty} \beta^i \exp\left[a_i + b_i\left(\rho(x_{t-1} - \mu)\right) + \xi_t\right] \equiv v_t\left(x_{t-1}, \xi_t\right) \end{aligned}$$

Denote $v_t(x_{t-1}, \xi_t)$ as $g(\xi_t|I_{t-1})$. Since $b_i > 0$ for all $i = 1, \cdots, \infty$ if $\rho > 0$, and $b_i < 0$ for all $i = 1, \cdots, \infty$ if $\rho < 0$. If $\rho = 0$, v_t is degenerate. So we exclude the case of $\rho = 0$. $g(\xi_t|I_{t-1})$ is a monotone increasing or decreasing function of ξ_t given I_{t-1} since it's a sum of monotone increasing or decreasing function. Let

$$\begin{aligned} r_t\left(x_{t-1}, \xi_t\right) &= \log\left(1 + g\left(\xi_t|I_{t-1}\right)\right) - \log\left(v_{t-1}(x_{t-1})\right) + (1-\rho) + \rho x_{t-1} + \xi_t \\ &\equiv G\left(\xi_t|I_{t-1}\right). \end{aligned}$$

If $b_i > 0$, $G(\xi_t|I_{t-1})$ is a monotone increasing function of ξ_t given I_{t-1}. It implies that there is an one-to-one transformation between ξ_t and $G(\xi_t|I_{t-1})$. The conditional probability density of r_t given I_{t-1} is

$$f_{r_t|I_{t-1}}\left(r_t|I_{t-1}\right) = \left.\frac{f_{\xi_t}\left(\xi_t\right)}{\left|\partial_{\xi_t} G\left(\xi_t|I_{t-1}\right)\right|}\right|_{\xi_t = G^{-1}\left(r_t|I_{t-1}\right)}.$$

$|\partial_{\xi_t} g(\xi_t|I_{t-1})| > 0$ since $g(\xi_t|I_{t-1})$ is monotone in ξ_t. $|\partial_{\xi_t} g(\xi_t|I_{t-1})| < \infty$ since by Assumption 2

$$\begin{aligned} &\partial_{\xi_t} g\left(\xi_t|I_{t-1}\right) = \sum_{i=1}^{\infty} \beta^i b_i \exp\left[a_i + b_i\left(\rho(x_{t-1} - \mu)\right) + \xi_t\right] \equiv \sum_{i=1}^{\infty} \tilde{z}_i \\ &\lim_{i\to\infty}\left(\tilde{z}_{i+1}/\tilde{z}_i\right) = \rho \exp\left(\alpha\mu\right) M_{\xi_t}\left(\theta\right) < 1. \end{aligned}$$

$0 < \partial_{\xi_t} G(\xi_t|I_{t-1}) < \infty$. Therefore, $0 < f_{r_t|I_{t-1}}(r_t|I_{t-1}) < \infty$ by Assumption 2. If $b_i < 0$, $\partial_{\xi_t} G(\xi_t|I_{t-1}) = \partial_{\xi_t} g(\xi_t|I_{t-1}) / (1 + g(\xi_t|I_{t-1})) + 1 = 0$ has only one solution for ξ_t given I_{t-1}, since $-\partial_{\xi_t} G(\xi_t|I_{t-1})$ is monotone decreasing in ξ_t and $1 + g(\xi_t|I_{t-1})$ is monotone increasing in ξ_t given I_{t-1}. It implies that there exists a partition B_1, B_2 such that for each t, there is an one-to-one transformation between $G_{B_k}(\xi_t|I_{t-1})$ and ξ_t on each B_k, $k = 1, 2$. Then, the conditional probability density of r_t given I_{t-1} is

$$f_{r_t|I_{t-1}}(r_t|I_{t-1}) = \sum_{k=1}^{2} \left. \frac{f_{\xi_t}(\xi_t)}{|\partial_{\xi_t} G_{B_k}(\xi_t|I_{t-1})|} \right|_{\xi_t = G_{B_k}^{-1}(r_t|I_{t-1})}.$$

$0 < f_{r_t|I_{t-1}}(r_t|I_{t-1}) < \infty$ then follows for $b_i < 0$. The joint conditional probability density of $r_t, \cdots, r_{t+n-1}$ given I_{t-1} is

$$\begin{aligned}
& f_{r_t,\ldots,r_{t+n-1}|I_{t-1}}(r_t, \cdots, r_{t+n-1}|I_{t-1}) \\
&= f_{r_t|I_{t-1}}(r_t|I_{t-1}) f_{r_{t+1}|r_t,I_{t-1}}(r_{t+1}|r_t, I_{t-1}) \cdots \\
& f_{r_{t+n-1}|r_{t+n-2},\ldots,r_t,I_{t-1}}(r_{t+n-1}|r_{t+n-2}, \cdots, r_t, I_{t-1}) \\
&= f_{r_t|I_{t-1}}(r_t|I_{t-1}) f_{r_{t+1}|I_t}(r_{t+1}|I_t) \cdots f_{r_{t+n-1}|I_{t+n-2}}(r_{t+n-1}|I_{t+n-2})
\end{aligned}$$

Since given I_{t-1}, r_t and x_t has one-to-one transformation on B_k, $k = 1, 2$, given r_t and I_{t-1} is the same as given I_t. The last equality then follows. Therefore, $0 < f_{r_t,\ldots,r_{t+n-1}|I_{t-1}}(r_t, \cdots, r_{t+n-1}|I_{t-1}) < \infty$. Consider the transformation of $(r_t, \ldots, r_{t+n-1})$ to $(U, U_1, \ldots, U_{n-1}) = \left(\sum_{j=0}^{n-1} r_{t+j}, r_{t+1}, \ldots, r_{t+n-1}\right)$. The joint probability density of $(U, U_1, \ldots, U_{n-1})$ given I_{t-1} is

$$\begin{aligned}
& f_{U,U_1,\ldots,U_{n-1}|I_{t-1}}(u, u_1, \ldots, u_{n-1}|I_{t-1}) \\
&= \left. \frac{f_{r_t,\ldots,r_{t+n-1}|I_{t-1}}(r_t, \ldots, r_{t+n-1}|I_{t-1})}{|J|} \right|_{r_t = u - \sum_{j=1}^{n-1} u_j, r_{t+1} = u_1, \ldots, r_{t+n-1} = u_{n-1}} \\
&= \left. f_{r_t,\ldots,r_{t+n-1}|I_{t-1}}(r_t, \ldots, r_{t+n-1}|I_{t-1}) \right|_{r_t = u - \sum_{j=1}^{n-1} u_j, r_{t+1} = u_1, \ldots, r_{t+n-1} = u_{n-1}}
\end{aligned}$$

Therefore, we have $0 < f_{U,U_1,\ldots,U_{n-1}|I_{t-1}}(u, u_1, \ldots, u_{n-1}|I_{t-1}) < \infty$.

Then, Lemma 7(i) is obvious. Lemma 7(ii) follows since

$$\begin{aligned} h_{\theta,t}(r_{t,n}|I_{t-1}) &= f_{U|I_{t-1}}(u|I_{t-1}) \\ &= \int f_{U,U_1,\ldots,U_{n-1}|I_{t-1}}(u, u_1, \ldots, u_{n-1}|I_{t-1})\, du_1 \cdots du_{n-1}, \end{aligned}$$

where $f_{U,U_1,\ldots,U_{n-1}|I_{t-1}}(u, u_1, \ldots, u_{n-1}|I_{t-1}) > 0$.

By the proposition that any function $f \in L^1(\omega, \mathcal{F}, \mu)$, then $|f| < \infty$. $h_{\theta,t}(r_{t,n}|I_{t-1}) < \infty$ since $\int f_{U|I_{t-1}}(u|I_{t-1})\, du = 1$.

By Assumption 1(iii), $f_{U,U_1,\ldots,U_{n-1}|I_{t-1}}(u, u_1, \ldots, u_{n-1}|I_{t-1})$ is continuously differentiable. From the mean value theorem, we have

$$|h_{\theta,t}(\lambda_1|I_{t-1}) - h_{\theta,t}(\lambda_2|I_{t-1})| = h'_{\theta,t}(c|I_{t-1})\, |\lambda_1 - \lambda_2|\, ,$$

where $c \in (\lambda_1, \lambda_2)$. If $h'_{\theta,t}(c|I_{t-1}) \leq L_0$, then Lemma 7(iv) holds.

■

Proof of Lemma 8: Applying the mean value theorem, we have

$$\left|q_{\theta,t}(\cdot, \delta_{\theta,n}) - q_{\theta,t}(\cdot, \delta^0_{\theta,n})\right| = \left|\nabla q_{\theta,t}(\cdot, \delta^*_{\theta,n})(\delta_{\theta,n} - \delta^0_{\theta,n})\right|,$$

where $\delta^*_{\theta,n} \in A$ and lies between $\delta_{\theta,n}$ and $\delta^0_{\theta,n}$.[3] Lemma 10 indicates that for fixed t and $\delta^*_{\theta,n} \in A$, the components of $\nabla q_{\theta,t}(\cdot, \delta^*_{\theta,n})$ are linearly independent of each other almost surely, which means that $\nabla q_{\theta,t}(\cdot, \delta^*_{\theta,n})(\delta_{\theta,n} - \delta^0_{\theta,n}) = 0$ if and only if $\delta_{\theta,n} - \delta^0_{\theta,n}$ is zero. If $\left\|\delta_{\theta,n} - \delta^0_{\theta,n}\right\| > \tau$ for every $\tau > 0$, then $\nabla q_{\theta,t}(\cdot, \delta^0_{\theta,n})(\delta_{\theta,n} - \delta^0_{\theta,n}) \neq 0$. Therefore, $\left|\nabla q_{\theta,t}(\cdot, \delta^0_{\theta,n})(\delta_{\theta,n} - \delta^0_{\theta,n})\right| > 0$ with positive probability. This implies that there exists $\delta_\tau > 0$, such that $P\left(\left|q_{\theta,t}(\cdot, \delta_{\theta,n}) - q_{\theta,t}(\cdot, \delta^0_{\theta,n})\right| > \delta_\tau\right) > 0$. ■

[3]Does $\beta_{\theta,n} \neq 0$ influence the use of the mean value theory?

Proof of Lemma 9: Q^0 is nonnegative definite. For any vector $p = (p_1, p_2, p_3, p_4)'$, we have

$$p'Q^0p = E\left[h_{\theta,t}\left(0|I_{t-1}\right)\left(\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p\right)'\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p\right]$$
$$= E\left[h_{\theta,t}\left(0|I_{t-1}\right)\left(\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p\right)^2\right] \geq 0.$$

Lemma 7 indicates that $h_{\theta,t}\left(0|I_{t-1}\right) > 0$. So, $p'Q^0p = 0$ if and only if $p\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right) = 0$ almost surely. Lemma 10 indicates that the components of $\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)$ are linearly independent almost surely, so there is no nontrival solution of p such that $p'Q^0p = 0$. Therefore, Q^0 is positive definite.

V^0 is nonnegative definite since

$$p'V^0p = E\left[\psi_\theta\left(\varepsilon_{\theta,t}\right)\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p\right]^2 \geq 0.$$

The equality holds if and only if $\psi_\theta\left(\varepsilon_{\theta,t}\right)\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p = 0$ almost surely. $\psi_\theta\left(\varepsilon_{\theta,t}\right) = \theta - 1_{\left\{\varepsilon_{\theta,t}<0\right\}}$ is nonzero, since $\psi_\theta\left(\varepsilon_{\theta,t}\right) = \theta$ or $\theta - 1$. Lemma 10 indicates that the components of $\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)$ are linearly independent almost surely, so there is no nontrival solution of p such that $\nabla q_{\theta,t}\left(\cdot,\delta^0_{\theta,n}\right)p = 0$ holds. Therefore, V^0 is positive definite. ■

Chapter 2

Dynamic Quantile Models for Risk Management

2.1 Introduction

Koenker and Bassett (1978) propose a regression quantile framework and establish the consistancy and asymptotic normality of the quantile regression estimators. The regression quantile model of Koenker and Bassett (1978) is a static quantile model. Engle and Manganelli (2004) introduce a conditional autoregressive value at risk (CAViaR) model, which is a dynamic quantile model. This model makes the calculation of conditional quantile and conditional value at risk possible. This paper also provides a test, called dynamic quantile (DQ) test, to evaluate the goodness of fit of estimated dynamic quantile process.

Other dynamic quantile models include the Quantile Autoregressive model (QAR) of Koenker and Xiao (2006), the Dynamic Additive Quantile (DAQ) model of Gouriéroux and Jasiak (2008) and the multi-quantile generalization of Engle and Manganelli's (2004) CaViaR approach to model conditional quantiles of White, Kim, and Manganelli (2008).

Ghysels, Plazzi, and Valkanov (2011) introduce a **MI**xed **DA**ta **S**ampling (MIDAS) quantile regression model, which address the conditional quantile of multiple horizon returns using single horizon returns(e.g. daily returns). Chen, Ghysels, and Wang (2010) introduce the class of models **H**igh Frequenc**Y** Data-**B**ased P**R**oject**I**on-**D**riven

(HYBRID) GARCH models, which addresses the issue of volatility forecasting involving forecast horizons of a different frequency. The HYBRID GARCH class of models allow us to write model multiple horizon models in a framework similar to GARCH(1,1). We adopt the same strategy for dynamic quantile models. That is, we introduce dynamic HYBRID quantile models that nest the CaViAR model of Engle and Manganelli (2004) and the MIDAS quantile models of Ghysels, Plazzi, and Valkanov (2011).

Sakata and White (1998) and Hall and Yao (2003) show that, for heavy-tailed errors, the asymptotic distributions of quasi-maximum likelihood parameter estimators in GARCH models are non-normal, and are particularly difficult to estimate directly using standard parametric methods. In such circumstances, dynamic quantile regression approaches might perform better than standard QMLE. We will show this by simulation in Section 2.5.

The conditional quantiles are typically not the direct object of interest. Instead, its key components, the conditional mean, conditional variance and the distribution are the prime focus. One may wonder how to obtain the predictive distribution of returns. Wu and Perloff (2005), Wu (2006) and Wu and Perloff (2007) proposed methods to fit densities to quantiles. Motivated by these methods, we propose a quantile distribution fits method to obtain conditional densities by matching the quantiles of a specific parametric family with the selected set of conditional quantiles.

This chapter is structured as follows. Section 2.2 describes the generic setup. Section 2.3 proposes models of financial risk based on dynamic quantile regressions. Section 2.4 introduces a density fitting approach to obtain conditional distributions of future returns based on matching conditional quantiles and parametric densities. 2.5 is the simulations of dynamic quantile regressions compared with conditional heteroskedasticity and quantile distribution fits for risk management. Section 2.6 concludes this chapter.

2.2 The Generic Setup

In this section, we describe the notations that will be used in the later sections.

Let us start with a location scale family. Let r_t be the portfolio return. We assume the return r_t follows

$$r_t = \mu_{t|t-1}(\theta_l^a) + \sqrt{\sigma_{t|t-1}^2(\theta_v^a)}\varepsilon_t \tag{2.1}$$

where $\mu_{t|t-1}(\theta_l^a)$ is conditional mean or conditional location using information $\Im_{t-1}$, $\sigma_{t|t-1}(\theta_v^a)$ is the conditional volatility using information $\Im_{t-1}$, and ε_t are i.i.d with $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = 1$, and density $F(\theta_d^a)$. Then the standardized return ε_t can be written as

$$\varepsilon_t(\theta^a) \equiv \frac{r_t - \mu_{t|t-1}(\theta_l^a)}{\sigma_{t|t-1}(\theta_v^a)} \tag{2.2}$$

where the parameter vector $\theta^a \equiv (\theta_l^{a\prime}, \theta_v^{a\prime}, \theta_d^{a\prime})'$ governs the location, scale and distribution of the standardized returns or returns.

Then the quantile function of the standardized return $\varepsilon_t(\theta^a)$ can be written as

$$Q^\varepsilon(p, \theta^a) = \inf\{\varepsilon \in R : p \leq F(\varepsilon, \theta_d^a)\} \tag{2.3}$$

where $0 < p < 1$ is a probability. Then the conditional quantile of return r_t can be written as

$$Q_t^r(p, \theta^a) = \mu_{t|t-1}(\theta_l^a) + Q^\varepsilon(p, \theta^a)\sigma_{t|t-1}(\theta_v^a) \tag{2.4}$$

The skewness and kurtosis of ε_t, if any, are not dynamic since ε_t are i.i.d. So the first two conditional moments, the conditional mean/location and conditional volatility, govern the dynamic of the conditional quantiles of r_t.

There are some evidence that the financial returns have some distributional predictable patterns that can not be fully captured by location-scale family in (2.1). Some literature shows that ε_t given by (2.2) have predictable patterns in skewness and kurtosis. These include Engle and Manganelli (2004), Kim and White (2004), Engle and Mistry (2007), White, Kim, and Manganelli (2008), (2010), Ghysels, Plazzi, and Valkanov (2011) and (2010b).

The bulk of the ARCH literature assumes that standardized returns normalized by conditional volatility is independent and identical distributed(i.i.d.). Francq and Zakoian (2004) have proved that quasi-maximum likelihood estimators(QMLE) for generalized autoregressive conditional heteroscedastic (GARCH) process and autoregressive moving-average(ARMA) GARCH process with i.i.d. innovations are consistent and asymptotically normal. To model higher order moments, one need extend the i.i.d assumptions on the innovations to some less restrictive assumptions. Escanciano (2009) has extended the consistency and asymptotic normality of the QMLE for pure GARCH process in Francq and Zakoian (2004) with i.i.d. innovations to martingale difference centered squared innovations. This extension is important since now the ARCH process allows for conditional skewness.

Now, let us consider the return r_t follows (2.1) where ε_t satisfies $E\left[\varepsilon_t|\Im_{t-1}\right] = 0$, $E[\varepsilon_t^2|\Im_{t-1}] = 1$ a.s., and has density $F(\theta_d^a)$. Note ε_t are not i.i.d. Assume the dependency of the quantile function of ε_t are governed by parameter θ^q. Then the dynamic quantile function of the standardized return can be written as

$$Q_t^{\varepsilon}(p, \theta^a, \theta^q) = \inf\left\{\varepsilon_t \in R : p \leq F(\varepsilon_t, \theta_d^a)\right\} \tag{2.5}$$

In conclusion, considering a location-scale model with relaxed assumption [1], we can study the dynamic quantile model $Q_t^{\varepsilon}(p, \theta^a, \theta^q)$ of the standardized return ε_t. We can also consider to model the conditional quantiles of return $Q_t^r(p, \theta^q)$ directly, where θ^q is the parameter determine the dynamic quantiles of return. This is a case beyond location-scale family. We can also further construct conditional mean/location, conditional volatility from the conditional quantiles of return $Q_t^r(p, \theta^q)$.

Here is an example of how to construct the predictive distribution [2] of return. Assume r_t is from a location-scale family, $\sigma_{t|t-1}(\theta_v^a)$ follows a GARCH(1,1), and $F(\theta_d^a)$ is zero mean unit variance Gaussian distribution. So the predictive distribution of return given $\Im_{t-1}$ is $r_t|\Im_{t-1} \sim N(\mu_{t|t-1}(\theta_l^a), \sigma_{t|t-1}(\theta_v^a))$. Now, we construct predictive distribution of r_t with conditional quantiles estimated through quantile models $Q_t^r(p, \theta^q)$. Define the interquartile range as

$$IQR_t^r(\theta^q) \equiv (Q_t^r(.75, \theta^q) - Q_t^r(.25, \theta^q)) \tag{2.6}$$

The predictive distribution of returns is $r_t|\Im_{t-1} \sim N(Q_t^r(.50, \theta^q), .549554 \times IQR_t^r(\theta^q)^2)$. .549554 is a constant using conditional quantiles to construct conditional volatility. If we need construct conditional skewness from conditional quantiles, we can adopt a robust coefficient of skewness proposed by Bowley. The conditional version of the measure of Bowley is as follows

$$Skew\,(r_t|\Im_{t-1}) = \frac{(Q_t^r(.75, \theta^q) - Q_t^r(.50, \theta^q)) - (Q_t^r(.50, \theta^q) - Q_t^r(.25, \theta^q))}{IQR_t^r(\theta^q)} \tag{2.7}$$

where $Q_t^r(.25, \theta^q)$, $Q_t^r(.50, \theta^q)$ and $Q_t^r(.75, \theta^q)$ are the 25th, 50th, and 75th conditional

[1] We can assume the normalized returns are a martingale difference sequence (see e.g. Escanciano (2009))

[2] i.e. conditional mean, conditional volatility, and conditional skewness, etc

quantiles of r_t.

For the cases that the conditional distribution can not be fully characterized by the first two or three moments, to obtain the predictive distribution of returns, we propose an Quantile Distribution Fits approach. Namely, we can use a parametric family to fit a conditional density via matching the quantiles of the parametric facility $q_t(p, \theta^d)$ with the selected set of conditional quantiles $Q_t^r(p, \theta^q)$ or $Q_t^\varepsilon(p, \theta^a, \theta^q)$ by the method of least squares.

2.3 Dynamic Quantile Models

Chen, Ghysels, and Wang (2010) introduce the class of models **H**igh Frequenc**Y** Data-**B**ased P**R**oject**I**on-**D**riven (HYBRID) GARCH models, which addresses the issue of volatility forecasting involving forecast horizons of a different frequency. Their HYBRID GARCH models can handle volatility forecasts for example over the next five business days with past daily data, or tomorrow's expected volatility while using intra-daily returns.

The HYBRID GARCH model(Chen, Ghysels, and Wang (2010)) has the following dynamics for volatility:

$$V_{\tau+1|\tau} = \omega + \alpha V_{\tau|\tau-1} + \beta H_\tau \tag{2.8}$$

where τ refers to a different time scale than t. When H_τ is simply a daily return we have the volatility dynamics of a standard daily GARCH(1,1), or H_τ a weekly return those of a standard weekly GARCH(1,1).

By further specify H_τ as

$$H_\tau \equiv H(\theta^H, \vec{r}_\tau) = \left[\sum_{j=1}^{m} \exp\left(\sum_{i=1}^{j} \left(\theta_0^H + \theta_1^H i/m + \theta_2^H i^2/m^2 \right) \right) r_{j,\tau}^2 \right] \tag{2.9}$$

where $\vec{r}_\tau = (r_{1,\tau}, r_{2,\tau}, \ldots, r_{m-1,\tau}, r_{m,\tau})^T$ is $\mathbb{R}^m-$valued random vector. The parameters to be estimated are $(\omega, \alpha, \beta, \theta_0^H, \theta_1^H, \theta_2^H)$ for the HYBRID GARCH model. We denote H_τ as given by 2.9 as exponential weights HYBRID GARCH model.

Ghysels, Plazzi, and Valkanov (2011) introduce a **MI**xed **DA**ta **S**ampling (MIDAS) quantile regression model, which addresses the conditional quantile of multiple horizon returns using single horizon returns(eg. daily returns). The MIDAS quantile regression model(Ghysels, Plazzi, and Valkanov (2011)) is described as follows.

$$Q_{\theta,t}\left(r_{t,n}; \delta_{\theta,n}\right) = \alpha_{\theta,n} + \beta_{\theta,n} Z_t\left(\kappa_{\theta,n}\right) \tag{2.10}$$

$$Z_t\left(\kappa_{\theta,n}\right) = \sum_{d=1}^{D} w_d\left(\kappa_{\theta,n}\right) x_{t-d} \tag{2.11}$$

where $\delta_{\theta,n} = \left(\alpha_{\theta,n}, \beta_{\theta,n}, \kappa_{\theta,n}\right)'$ are unknown parameters to estimate. Following Ghysels, Santa-Clara, and Valkanov (2006), we can specify $\omega_d\left(\kappa_{\theta,n}\right)$ as

$$\omega_d\left(\kappa_{\theta,n}\right) = \frac{f\left(\frac{d-1/2}{D}, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)}{\sum_{m=1}^{D} f\left(\frac{m-1/2}{D}, \kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)}, \tag{2.12}$$

where $\kappa_{\theta,n} = \left(\kappa_{1,\theta,n}, \kappa_{2,\theta,n}\right)$ is a 2-dimensional row vector that reduces the number of weights for lag coefficient to estimate from D to 2, $f\left(z, a, b\right) = z^{a-1}\left(1-z\right)^{b-1}/\beta\left(a, b\right)$, $\beta\left(a, b\right) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, and Γ is Gamma function. We denote Z_t as given by 2.12 as beta weights MIDAS Quantile model.

Engle and Manganelli (2004) introduce Conditional Autoregressive Value at Risk (CAViaR) model, which is a quantile regression model specified as follows.

$$Q_t\left(\beta\right) = \beta_0 + \sum_{i=1}^{q} \beta_i Q_{t-i}\left(\beta\right) + \sum_{j=1}^{r} \beta_j l\left(\mathbf{x_{t-j}}\right) \tag{2.13}$$

where $p = q + r + 1$ is the dimension of β and l is a function of a finite number of lagged values of observations.

The HYBRID GARCH class of models allowed us to propose multiple horizon models in a framework similar to GARCH(1,1). We adopt the same strategy for dynamic quantile models. That is, we introduce dynamic HYBRID quantile models that nest (1) the CaViAR model of Engle and Manganelli (2004) and (2) the MIDAS quantile models of Ghysels, Plazzi, and Valkanov (2011).

We characterize a HYBRID quantile regression in a similar way to HYBRID GARCH - where the conditional quantile pertains to multiple horizon returns and the regressors are higher frequency returns - as follows:

$$\begin{aligned} Q^r_\tau(p,\theta^q) &= \omega + \alpha Q^r_{\tau-1}(p,\theta^q) + \beta H^Q_\tau && (2.14) \\ H^Q_\tau &= \sum_{j=0}^{m-1} w_j\left(\kappa\right) x_{j,\tau} && (2.15) \end{aligned}$$

when the HYBRID process driving the quantile is a same frequency absolute return we recover the CaViAR model, and when $\alpha = 0$ we recover the MIDAS quantile. There are several benefits from using the HYBRID and MIDAS quantile specification (2.14)-(2.15) rather than other conditional quantile models, such as Engle and Manganelli (2004) and White, Kim, and Manganelli (2008). We follow Engle and Manganelli (2004), who find that absolute returns successfully capture time variation in the conditional distribution of returns, and use absolute daily or intra-daily returns as the conditioning variable in (2.15). Alternative specifications with squared returns will be considered also.

To test the validity of the forecast model of CAViaR, Engle and Manganelli (2004) propose a new test, in-sample DQ test, which is used for model selection. The test is defined as follows.

$$DQ_{IS} \equiv \frac{\hat{Hit}'\left(\hat{\beta}\right)\hat{X}\left(\hat{\beta}\right)\left(\hat{\mathbf{M}}_T\hat{\mathbf{M}}'_T\right)^{-1}\hat{X}'\left(\hat{\beta}\right)\hat{Hit}'\left(\hat{\beta}\right)}{\theta\left(1-\theta\right)} \overset{d}{\sim} \chi^2_q \; as \; T \to \infty \qquad (2.16)$$

where Hit is defined as follows.

$$Hit(\beta^{\mathbf{0}}) \equiv I\left(r_t < Q_t\left(\beta_{\mathbf{0}}\right)\right) - \theta \tag{2.17}$$

Further definitions of $\mathbf{X}\left(\hat{\beta}\right)$ and $\hat{\mathbf{M}}_T$ can be found in Engle and Manganelli (2004).

We use S&P 500 daily returns ranging from 1982 to 2011 to test our HYBRID quantile models. We will estimate a generic of HYBRID quantile models with both exponential weight(2.9) and beta weights(2.12). The choice of x in (2.15) we use are $|r|$, r^2, r^3 and r. We estimate 1% and 5% weekly VaRs(horizon 5) using non-overlapping daily returns with lag 5.

Table 2.1 shows the estimated parameters obtained from HYBRID quantile models and MIDAS quantile models for 5% VaRs. Both Hit and DQ test p values are for in-sample tests. Hit in percent is the percentage of times that the VaR is exceeded. As indicated by Hit, the precision of all the models are good. Most of quantile models are not rejected at 5% confidence interval by DQ tests for exponential weights except three of the MIDAS quantile models. For beta weights, HYBRID quantile models are also prefered by DQ in-sample test.

Table 2.2 shows the estimated parameters obtained from HYBRID quantile models and MIDAS quantile models for 1% VaRs. The models perform similarly by looking at in-sample Hit and DQ tests for 1% VaRs.

Figure 2.1 shows the 5% through 95% multiple horizon quantiles (horizon 5) obtained using HYBRID quantile regression method and MIDAS quantile regression method using daily returns with lag 5. As expected, with the lag term of quantile included in the HYBRID quantile regression, the quantiles obtained are more smoother than the quantiles obtained from MIDAS quantiles.

2.4 Quantile Distribution Fits

Wu and Perloff (2005), Wu (2006) and Wu and Perloff (2007) fits densities to quantiles. This is an interesting aspect if we have several conditional quantiles and we want to use them to find the conditional density of either returns or standard returns by fitting quantiles to a density. We call this method Quantile Distribution Fits.

Assume we have conditional quantiles $Q_t^r(p, \theta^q)$ for a selection of p-values and determined by a parameter vector θ^q for return r at time t. The $Q_t^r(p, \theta^q)$ can be obtained by quantile regression method like CAViaR, MIDAS Quantile regression, and HYBRID Quantile regression. Then the conditional distribution of r at time t can be found by solving

$$\min_{\theta^d} \frac{1}{N} \sum_{p=1}^{N} \left[Q_t^r(p, \theta^q) - q_t(p, \theta_d)\right]^2, \ \forall t \in \{1, ..., T\} \tag{2.18}$$

where θ^d is the parameters to be estimated, N is the number of quantiles used in finding conditional distribution, and $q_t(p, \theta_d)$ is the quantile function of selected distribution.

For the choice of $q_t(p, \theta_d)$, we can pick a rich family of distributions, like the Generalized Hyperbolic (GH) class which is characterized by five parameters. When further narrowed down to subclasses of four-, three-, or two-parameter distributions, yields widely used distributions such as the normal inverse Gaussian distribution, the hyperbolic distribution, the variance gamma distribution, the generalized skewed t distribution, the student t distribution, the gamma distribution, the Cauchy distribution, the normal distribution, etc. We can also use extreme value distributions like Generalized Extreme Value (GEV) distribution and Generalized Pareto (GP) distribution.

For the choice of N, we can in principle fit as many quantiles as we want. More quantiles means better distributional fit, but they may start crossing. The more quantiles we use, the issue of crossing becomes more acute and then there is also the issue of too many moment conditions, which creates singularities.

By having the conditional distribution, we can further obtain Expected Shortfall (ES), an alternative measure of risk proposed by Artzner, Delbaen, Eber, and Heath (1997). The Expected Shortfall is the expected value of r when the threshold (i.e. VaR) has been exceeded. It can be calculated by integral over the quantile function $q_t(p, \theta_d)$ in our case. The αth Expected Shortfall is defined as follows

$$ES_t^\alpha = E_t\left(r_t | r_t < q_t(\alpha, \theta_d)\right) = \frac{1}{\alpha}\int_0^\alpha q_t(\gamma, \theta_d) d\gamma \tag{2.19}$$

where $0 < \alpha < 1$.

We would like to compare the Expected Shortfall obtained using the fitted parameters of quantile distribution fits with the regression based Expected Shortfall for CaViaR or other quantile models(Manganelli and Engle (2001)). The regression based Expected Shortfall is defined as follows

$$r_t = \delta Q_t^r(p, \theta^q) + \eta_t,\ \ r_t < Q_t^r(\alpha, \theta^q) \tag{2.20}$$

$$\hat{ES}_t^\alpha = \hat{E}_t^\alpha\left(r_t | r_t < Q_t^r(\alpha, \theta^q)\right) = \hat{\delta} Q_t^r(\alpha, \theta^q) \tag{2.21}$$

We start fitting generalized extreme value(GEV) distribution to quantiles of return by minimizing the sum of squared distances of quantiles given by (2.18). The preliminary results are shown in Figure 2.2. The quantiles used in this figure were 10%, 20%, 30%, and 40% quantiles obtained by CAViaR SAV model using daily return. There are three parameters to be estiamted(location, scale and shape). The quantiles obtained by quantile distribution fits and CAViaR are generally on top of each other. The smaller the quantiles, the more discrepancy between quantiles obtained by two methods. Quantiles obtained by quantile distribution fits tend to be smaller for lower quantiles.

The results for comparison of Expected Shortfall using conditional distribution from quantile distribution fits and regression based Expected Shortfall are shown in Figure 2.3. The larger discrepancy for 1% ES may be caused by the smaller sample size in the regression.

We also test other distributions, including generalized pareto(GP) distribution. In general, quantile distribution fits with GEV performs better than with GP. Also, quantile distribution fits with t, skew t, and generalized hyperbolic distribution fails sometimes due to a lack of analytic quantile functions. We also use other quantiles like 25%, 50%, and 75% quantiles, and the results are worse than using 10%, 20%, 30%, and 40% quantiles.

2.5 Simulation

In Section 2.5.1, we present results to compare the simulation results to compare conditional heteroskedasticity and quantiles.

Simulation of Conditional Heteroskedasticity versus Quantils

This section covers an extensive Monte Carlo simulation to compare conditional heteroskedasticity and quantiles. We first describe the conditional heteroskedasticity and quantiles models we use in this section.

We consider the conditional volatility as GARCH(1,1)

$$r_t = \sigma_t \varepsilon_t \tag{2.22}$$

$$\sigma_t^2 = \omega_0 + \alpha_0 r_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \tag{2.23}$$

where $E\left[\varepsilon_t | \mathcal{F}_{t-1}\right] = 0$, and $E\left[\varepsilon_t^2 | \mathcal{F}_{t-1}\right] = 1$. By specifying the density of ε_t, we define seven GARCH type models.

If $\varepsilon_t \sim N(0,1)$, the model is Gaussian GARCH(1,1) and we denoted it as NOR. The parameters to be estimated for this model is $\theta = (\omega_0, \alpha_0, \beta_0)$.

If ε_t is Student's t-distribution which has the probability density function given by

$$f(t|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{2.24}$$

where $\nu > 2$ is the number of degree of freedom and Γ is the Gamma Function. We denote this Student's t GARCH model as STDT. The parameters to be estimated for this model is $\theta = (\omega_0, \alpha_0, \beta_0, \nu)$.

If ε_t is Skew t-distribution proposed by Hansen (1994) which has the probability density function given by

$$g(z|\nu,\lambda) = bc\left(1+\frac{1}{\nu-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{(-(\nu+1)/2)}, \; z < -a/b \tag{2.25}$$

$$= bc\left(1+\frac{1}{\nu-2}\left(\frac{bz+a}{1+\lambda}\right)^2\right)^{(-(\nu+1)/2)}, \; z \geq -a/b \tag{2.26}$$

where $\nu > 2$, $-1 < \lambda < 1$, and

$$a = 4\lambda c\frac{\nu-2}{\nu-1}$$

$$b^2 = 1+3\lambda^2 - a^2$$

$$c = \frac{\Gamma\frac{\nu+1}{2}}{\sqrt{\pi(\nu-1)}\Gamma(\nu/2)}.$$

To ensure the mean and variance of ε_t to be zero, a, b, and c must satisfy

$$E[Z] = a = 0$$

$$E\left[Z^2\right] = b^2 + a^2 = 1.$$

We denote this SKWE T GARCH model as SKEWT. The parameters to be estimated for this model is $\theta = (\omega_0, \alpha_0, \beta_0, \nu, \lambda)$. Note there are only one free parameter λ to be estimated, and it is the skewness parameter of this density.If $\lambda > 0$, the density is positively skewed and vice versa.

If ε_t is Generalized Hyperbolic Skew Student's t-distribution proposed by Aas and Haff (2006) which has the probability density function given by

$$f(x|\beta,\nu,\mu,\delta) = \frac{2^{\frac{1-\nu}{2}}\delta^{\nu}\left|\beta\right|^{\frac{\nu+1}{2}} K_{\frac{\nu+1}{2}}\left(\sqrt{\beta^2\left(\delta^2+(x-\mu)^2\right)}\right)\exp\left(\beta\left(x-\mu\right)\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}\left(\sqrt{\delta^2+(x-\mu)^2}\right)^{\frac{\nu+1}{2}}},\ \beta \neq 0 \tag{2.27}$$

$$= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\delta\Gamma\left(\frac{\nu}{2}\right)}\left[1+\frac{(x-\mu)^2}{\delta^2}\right]^{-(\nu+1)/2},\ \beta = 0 \tag{2.28}$$

where $\nu > 4$ to ensure finite variance. To ensure the mean and variance of ε_t to be zero, the parameters must satisfy

$$E\left[X\right] = \mu + \frac{\beta\delta^2}{\nu-2} = 0$$
$$Var\left[X\right] = \frac{2\beta^2\delta^4}{(\nu-2)^2(\nu-4)} + \frac{\delta^2}{\nu-2} = 1$$

We denote this Generalized Hyperbolic Skew t GARCH model as GHST. The parameters to be estimated for this model is $\theta = (\omega_0, \alpha_0, \beta_0, \beta, \nu, \mu, \delta)$.

The skewness of the above density is

$$skew\left[X\right] = \frac{2\left(\nu-4\right)^{1/2}\beta\delta}{\left[2\beta^2\delta^2+(\nu-2)(\nu-4)\right]^{3/2}}\left[3\left(\nu-2\right)+\frac{8\beta^2\delta^2}{\nu-6}\right]. \tag{2.29}$$

It is time-invariant. To generate time-varying skewness in the simulation, we also consider two Generalized Hyperbolic Skew t GARCH models with either ν or β follow a AR(1) process.

$$\nu_t = c + \phi\nu_{t-1} + \epsilon_t \tag{2.30}$$

$$\beta_t == c + \phi\beta_{t-1} + \epsilon_t \tag{2.31}$$

where ϵ_t is white noise with variance k. We denote the Generalized Hyperbolic Skew t GARCH with time-varying β model as GHYP1 and the Generalized Hyperbolic Skew t GARCH with time-varying ν model as GHYP2. The parameters for this model is $\theta = (\omega_0, \alpha_0, \beta_0, \beta, \nu, \mu, \delta, c, \phi, k)$. The last three parameters are determined without estimation for both GHYP1 and GHYP2.

The last GARCH type model we consider is the model that ε_t follows mixed normal distribution with two components. We denote this model as MIXNOR. The parameters to be estimated for this model is $\theta = (\omega_0, \alpha_0, \lambda_1, \lambda_2, \mu_1, \mu_2, \sigma_1, \sigma_2)$. These parameters must satisfy conditions such that $\lambda_1 + \lambda_2 = 1$, $E\left(\varepsilon_t|\mathcal{F}_{t-1}\right) = 0$, and $E\left(\varepsilon_t^2|\mathcal{F}_{t-1}\right) = 1$.

The single horizon quantile models we consider here are four CAViaR models proposed by Engle and Manganelli (2004). Let r_t be the return, and q_t be the θth quantile of r_t. The symmetric Absolute Value CAViaR model, denoted as SAV, is

$$q_t\left(\boldsymbol{\beta}\right) = \beta_1 + \beta_2 q_{t-1}\left(\boldsymbol{\beta}\right) + \beta_3\left|r_{t-1}\right|. \tag{2.32}$$

The Symmetric Square Value CAViaR model, denoted as SSV, is

$$q_t\left(\boldsymbol{\beta}\right) = \beta_1 + \beta_2 q_{t-1}\left(\boldsymbol{\beta}\right) + \beta_3 r_{t-1}^2. \tag{2.33}$$

The Asymmetric Slope CAViaR model, denoted as AS, is

$$q_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 q_{t-1}(\boldsymbol{\beta}) + \beta_3 (r_{t-1})^+ + \beta_4 (r_{t-1})^- . \tag{2.34}$$

The Adaptive CAViaR model, denoted as AD, is

$$q_t(\beta_1) = q_{t-1}(\beta_1) + \beta_1 \left\{ \left[1 + \exp\left(G[y_{t-1} - q_{t-1}(\beta_1)]^{-1} - \theta \right) \right] \right\}, \ G = 10. \tag{2.35}$$

Table 2.4 provides a summary of notations and descriptions of these models used in the simulation and estimation.

We simulate data using seven different data generating processes (i.e. NOR, STDT, SKEWT, GHYP, GHYP1, GHYP2, and MIXNOR). For the data generating processes NOR, STDT, SKEWT, GHYP and MIXNOR, the parameters used in the simulations are obtained by estimating 1982-2011 S&P 500 returns using the models accordingly. For GHYP1 and GHYP2, we use time-varying β and ν generated by AR(1) processes, respectively, while other parameters remain the same as GHYP. For each data generating process, we simulate 1000 samples with length 2500.

Table 2.4 shows all the parameter choices used in the simulation. They are obtained by estimating 1982-2011 S&P 500 daily, weekly, and biweekly returns using the models accordingly. The last column is log likelihood obtained through the estimations. For daily data, STDT model is the best model by looking at this criteria. For weekly and biweekly data, MIXNOR and GHYST provide the best estimation results, respectively.

For each sample, we estimate conditional heteroskedasticity models(NOR, STDT, SKEWT, GHYP, and MIXNOR) and CaViAR models(5%, 25%, and 75% quantiles). The performances of model estimations are evaluated through the estimates of $\hat{\sigma}_t$ and 5% VaR. Our purposes are to compare the conditional volatility and conditional Value at risk estimated through GARCH type models and quantile models. This raises the questions what are the true and estimated conditional Value at risk from GARCH type models, and how to find out the conditional volatility from the quantile models.

For CAViaR models, the $\hat{\sigma}_t^2$ is estimated through $c \times \hat{IQR}^2$, where c is a parameter estimated through the interquartile range of each DGP[3] and $\hat{IQR}$ is the estimates of interquartile range. For conditional heteroskedasticity models, the 5% VaR is estimated through $q_{5\%}^{true}\sigma_t^{true}$, where $q_{5\%}^{true}$ is the 5% quantile of each DGP.

The measures we use to compare $\hat{\sigma}_t$ are $QLIKE$ and $MSEprop$ proposed by Patton (2011). The definitions are as follows.

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T}\left(\log\frac{h_t}{\hat{\sigma}_t^2} + \frac{\hat{\sigma}_t^2}{h_t} - 1\right), \tag{2.36}$$

$$MSEprop = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{\sigma}_t^2}{h_t} - 1\right)^2, \tag{2.37}$$

and $h_t = (\sigma_t^{true})^2$. where $QLIKE$ is normalized to yield zero when the estimated volatility is equal to the true volatility. A smaller value of $QLIKE$ means better estimation. We compare the estimates of 5% VaR using Mean squared error.

The results of comparisons are shown in Table 2.5 - Table 2.7.

Table 2.5 shows the comparison of σ_t using $QLIKE$. For the simulation with data generating process NOR, the CaViaR quantile models SAV and AS perform comparably to the true model NOR. For data generating process STKEWT, the CaViaR quantile model SAV performs comparably to the true model SKEWT. GARCH type model NOR and CaViaR model AS perform similarly and slightly worse than the true model SKEWT. For data generating process GHST, the true model performs the best, then followed by other GARCH type models. In this case, the CaViaR quantile models

[3]For example, for GARCH-Normal, $c = .549554$.

do not show advantage over the GARCH type models. But for data generating process GHYP2, the CaViaR quantile models SAV performs comparably with estimated through GHYP and performs better than other GARCH type models. For data generating process MIXNOR, CaViaR quantile model SAV performs better than NOR, STDT, and GHST, and worse than SKEWT and the true model MIXNOR. Overall, CaViaR model SAV performs consistently very well for a variety of data generating process.

Table 2.6 shows the comparison of σ_t using $MSEprop$. For data generating process NOR, SAV performs similarly to NOR by looking $MSEprop$. For data generating process STDT, CaViaR quantile models SAV, SSV and AS perform even better than the true model STDT. For data generating process SKEWT, the CaViaR model SAV and AS perform better than the true model SKEWT. For data generating process GHST, the true model performs the best, then followed by other GARCH type models. In this case, the CaViaR quantile models do not show advantage over the GARCH type models as using the measure of $QLIKE$. For data generating process MIXNOR, CaViaR quantile model SAV performs the best. Overall, using $MSEprop$ as criteria, CaViaR quatile models shows even more advantages than GARCH type models compared with using $QLIKE$.

In conclusion, for estimation of $\hat{\sigma}_t$, CAViaR Models (SAV, SSV, AS) are better than GARCH type models when there are fat tail, skewness or time-varying skewness in the data.

Table 2.7 shows the comparison of VaR using MSE. And the findings can be summarized as follows. For estimation of VaR, some of the GARCH type models are better than CaViaR Models. This makes sense since the estimation of $q_{5\%}$ is less accurate than say the estimations of $q_{25\%}$ and $q_{75\%}$ for skewness measures.

2.6 Conclusion

We introduce a generic of HYBRID quantile regression models and use the measure of in-sample *Hit* and *DQ* tests(Manganelli and Engle (2001)) to check the performance of our models compared with MIDAS quantile regression models. For the estimation of 5% VaRs, the HYBRID quantile regression models are prefered. For 1% VaRs, there two types of models provide similar results.

We propose a method to find conditional distributions based on quantile regressions called Quantile Distribution Fits. This method allows us to calculate Expected Shortfall, and other properties, which is very useful for risk management. We compare the results of quantiles/Value at Risk by quantile regressions and quantile distribution fits. We also study the expected shortfall using conditional distribution obtained by quantile distribution fits with the regression based expected shortfall for quantiles regressions. The results suggest that Quantile Distribution Fits is a very promising alternative method for risk management.

For estimation of $\hat{\sigma}_t$, CAViaR Models (SAV, SSV, AS) are better than GARCH type models when there are fat tail, skewness or time-varying skewness in the data. For estimation of VaR, some of the GARCH type models are superior than CaViaR Models. This may arise from the fact that the estimation of $q_{5\%}$ is less accurate than say the estimations of $q_{25\%}$ and $q_{75\%}$ for skewness measures.

2.7 Tables and Figures

This section contains tables and figures for this chapter.

Table 2.1: Hybrid quantiles and MIDAS quantiles for 5% VaR

Model	HYBRID				MIDAS			
x	$\lvert r\rvert$	r^2	r	r^3	$\lvert r\rvert$	r^2	r	r^3
				Panel I: Exponential Weights				
ω	-0.2255	-0.5661	-0.6124	-0.9906	-1.8101	-2.8321	-3.8394	-3.4571
α	0.7201	0.7692	0.8408	0.6911				
β	-1.0231	-0.1710	1.0062	0.0407	-2.0661	-0.4581	0.8005	0.0466
κ_1	82.4187	18.8972	1.3519	223.6761	58.4813	4.2246	335.1269	239.4319
κ_2	-11.5922	-2.4032	-0.1867	-31.5316	-6.5666	-0.4442	-47.6295	-29.9681
Hit (%)	4.9366	5.0033	5.0033	5.0033	5.0033	5.0033	5.0033	4.9366
DQ p values	0.9370	0.8868	0.5496	0.8883	0.0172	0.9630	0.0000	0.0428
				Panel II: Beta Weights				
ω	-0.2018	-0.5769	-0.8949	-0.7841	-1.9384	-2.8254	-3.8074	-3.4578
α	0.7153	0.7692	0.7559	0.7565				
β	-1.0891	-0.1648	0.8543	0.0290	-1.8649	-0.4500	0.7887	0.0466
κ_1	70.3929	62.6558	10.5647	53.9638	152.6235	221.1039	21.8558	128.1018
κ_2	44.9371	37.6604	4.4327	29.8219	1.8488	1.8442	10.8169	3.2954
Hit (%)	5.0033	4.9366	4.9366	5.0033	5.0700	5.0033	5.0033	5.0033
DQ p values		0.9438	0.0965	0.5482			0.0000	

Table 2.2: Hybrid quantiles and MIDAS quantiles for 1% VaR

Model	HYBRID				MIDAS			
x	$\|r\|$	r^2	r	r^3	$\|r\|$	r^2	r	r^3
Panel I: Exponential Weights								
ω	-0.6471	-1.1271	-1.2982	-1.1205	-4.3155	-5.0566	-7.0351	-6.2308
α	0.7436	0.7436	0.8032	0.8200				
β	-0.9795	-0.2721	2.1812	0.0406	-2.1111	-0.7087	1.6703	0.0357
κ_1	16.7011	51.7878	0.0052	283.2151	58.7313	58.4610	6.8645	197.2404
κ_2	-2.0263	-6.0263	0.0242	-31.5316	-6.5666	-6.5666	-0.8859	-22.2962
Hit (%)	1.0007	1.0007	0.9340	1.0007	1.0007	0.9340	1.0007	1.0007
DQ p values	0.7737	0.9696	0.1539	0.8734	0.9495	0.9851	0.4555	
Panel II: Beta Weights								
ω	-0.6768	-1.1360	-1.6637	-2.2420	-4.2135	-5.0549	-6.6011	-6.2305
α	0.7436	0.7436	0.7391	0.6356				
β	-0.9266	-0.2619	0.9580	0.0310	-2.3949	-0.7095	1.2658	0.0346
κ_1	85.6395	134.1690	11.1725	64.2875	152.6235	192.1371	5.5065	160.3267
κ_2	53.1661	26.4930	4.7366	36.2934	1.8260	1.8402	1.7065	9.5178
Hit (%)	1.0007	1.0007	0.9340	1.0007	1.0007	1.0007	0.9340	1.0007
DQ p values	0.8524	0.9702	0.9959	0.9279			0.9724	

Table 2.3: Summary of Model Specifications

Model	Notation	Description
1	NOR	Gaussian GARCH
2	STDT	TGARCH
3	SKEWT	Skew T GARCH(Hansen (1994))
4	GHST	Generalized Hyperbolic Skew T GARCH(Aas and Haff (2006))
5	GHST1	Generalized Hyperbolic Skew T GARCH with Time Varying β (Aas and Haff (2006))
6	GHST2	Generalized Hyperbolic Skew T GARCH with Time Varying ν (Aas and Haff (2006))
7	MN(3,3)	Mixed Normal GARCH with 3 component densities and 3 GARCH process(Haas, Mittnik, and Paolella (2004))
8	MN	Mixed Normal GARCH
9	SAV	CAViaR: Symmetric Absolute Value
		$q_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 q_{t-1}(\boldsymbol{\beta}) + \beta_3 \lvert y_{t-1} \rvert$
10	SSV	CAViaR: Symmetric Square Value
		$q_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 q_{t-1}(\boldsymbol{\beta}) + \beta_3 y_{t-1}^2$
11	AS	CAViaR: Asymmetric Slop
		$q_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 q_{t-1}(\boldsymbol{\beta}) + \beta_3 (y_{t-1})^+ + \beta_4 (y_{t-1})^-$
12	AD	CAViaR: Adaptive
		$q_t(\beta_1) = q_{t-1}(\beta_1) + \beta_1 \left\{\left[1 + \exp\left(G[y_{t-1} - q_{t-1}(\beta_1)]^{-1} - \theta\right)\right]\right\}$, $G = 10$

Table 2.4: Summary of Parameters in Simulation Study

Model	**Parameters**	**LL**
NOR	$(\omega_0, \alpha_0, \beta_0)$	
daily	0.0133 0.0798 0.9115	-10295
weekly	0.1471 0.1337 0.8465	-3359
biweekly	0.1523 0.1016 0.8984	-1963
STDT	$(\omega_0, \alpha_0, \beta_0, \nu)$	
daily	0.0070 0.0571 0.9381 6.2893	-10057
weekly	0.1039 0.0935 0.8893 8.8307	-3336
biweekly	0.3028 0.1021 0.8693 5.5406	-1900
SKEWT	$(\omega_0, \alpha_0, \beta_0, \nu, \lambda)$	
daily	0.1004 0.1401 0.7683 21.5589 -0.0417	-10281
weekly	0.1080 0.0967 0.8897 8.4570 -0.1834	-3324
biweekly	0.5067 0.1558 0.8269 4.5074 -0.2634	-1885
GHST	$(\omega_0, \alpha_0, \beta_0, \beta, \nu, \mu, \delta)$	
daily	0.0000 0.1106 0.8894 -0.2681 13.1669 0.2641 3.3162	-10166
weekly	0.1014 0.0942 0.8949 -0.5415 8.9402 0.4890 2.5036	-3324
biweekly	0.5999 0.1706 0.8294 -0.5198 5.0646 0.3793 1.4954	-1886
GHST1	$(\omega_0, \alpha_0, \beta_0, \beta, \nu, \mu, \delta, c, \phi, k)$	
daily	0.0000 0.1106 0.8894 -0.2681 13.1669 0.2641 3.3162 -0.0330 0.8000 0.0260	
weekly	0.1014 0.0942 0.8949 -0.5415 8.9402 0.4890 2.5036 -0.0330 0.8000 0.0260	
biweekly	0.5999 0.1706 0.8294 -0.5198 5.0646 0.3793 1.4954 -0.0330 0.8000 0.0260	
GHST2	$(\omega_0, \alpha_0, \beta_0, \beta, \nu, \mu, \delta, c, \phi, k)$	
daily	0.0000 0.1106 0.8894 -0.2681 13.1669 0.2641 3.3162 -0.2305 0.8000 0.4000	
weekly	0.1014 0.0942 0.8949 -0.5415 8.9402 0.4890 2.5036 -0.2305 0.8000 0.4000	
biweekly	0.5999 0.1706 0.8294 -0.5198 5.0646 0.3793 1.4954 -0.2305 0.8000 0.4000	
MIXNOR	$(\omega_0, \alpha_0, \beta_0, \lambda_1, \lambda_2, \mu_1, \mu_2, \sigma_1, \sigma_2)$	
daily	0.0084 0.0650 0.9312 0.9322 0.0678 0.0469 -0.6450 0.8718 1.9627	-10082
weekly	0.0984 0.0920 0.8968 0.9057 0.0943 0.0858 -0.8246 0.8500 1.7074	-3323
biweekly	0.4242 0.1455 0.8545 0.9474 0.0526 0.0849 -1.5294 0.7565 2.4971	-1893

Table 2.5: Comparison of σ_t using QLIKE

The entries are summary statistics(mean on the top and std on the bottom) of QLIKE to compare the simulation results of σ_t, where $QLIKE = \frac{1}{T}\sum_{t=1}^{T}\left(\log\frac{h_t}{\hat{\sigma}_t^2} + \frac{\hat{\sigma}_t^2}{h_t} - 1\right)$, and $h_t = (\sigma_t^{true})^2$. Simulations are conducted using 2500 observations across 1000 trials. The row names are simulation models, and the column names are estimation models. NOR is Gaussian GARCH model. STDT is TGARCH model. SKEWT is Skew T GARCH model(Hansen (1994)). GHST is Generalized Hyperbolic Skew T GARCH model(Aas and Haff (2006)). GHST1 is Generalized Hyperbolic Skew T GARCH model with Time Varying β. GHYT2 is Generalized Hyperbolic Skew T GARCH model with Time Varying ν. MIXNOR is a mixed normal GARCH model where the error is a mixture of two normal distributions. SAV is Symmetric Absolute Value CAViaR model(Engle and Manganelli (2004)) for quantiles. SSV is Symmetric Square Value CAViaR model. AS is Asymmetric Slop CAViaR model. AD is Adaptive CAViaR model.

	Estimation								
Simulation	NOR	STDT	SKEWT	GHST	MIXNOR	SAV	SSV	AS	AD
NOR	0.0025					0.0080	0.0164	0.0099	0.0659
	0.0044					0.0039	0.0114	0.0058	0.0136
STDT	0.0058	0.0069				0.0105	0.0228	0.0128	0.0716
	0.0095	0.0306				0.0051	0.0217	0.0072	0.0149
SKEWT	0.0062	0.0060	0.0042			0.0104	0.0237	0.0123	0.0741
	0.0118	0.0159	0.0082			0.0045	0.0196	0.0058	0.0155
GHST	0.0065	0.0063	0.0047	0.0036		0.0631	0.0833	0.0657	0.1701
	0.0099	0.0148	0.0127	0.0029		0.0180	0.0377	0.0190	0.0292
GHST1	0.0124	0.0133	0.0059	0.0050		0.0358	0.0556	0.0380	0.1428
	0.0473	0.1759	0.0090	0.0058		0.0133	0.0447	0.0139	0.0299
GHST2	0.0267	0.0162	0.0172	0.0110		0.0147	0.0192	0.0168	0.0853
	0.0744	0.0125	0.0111	0.0096		0.0071	0.0214	0.0084	0.0220
MIXNOR	0.0090	0.0082	0.0059	0.0803	0.0047	0.0113	0.0331	0.0146	0.0762
	0.0122	0.0123	0.0072	0.2141	0.0046	0.0035	0.0285	0.0075	0.0087

Table 2.6: Comparison of σ_t using MSEprop

The entries are summary statistics(mean on the top and std on the bottom) of MSEprop to compare the simulation results of σ_t, where $MSEprop = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{\sigma}_t^2}{h_t} - 1\right)^2$. Simulations are conducted using 2500 observations across 1000 trials. The row names are simulation models, and the column names are estimation models. NOR is Gaussian GARCH model. STDT is TGARCH model. SKEWT is Skew T GARCH model(Hansen (1994)). GHST is Generalized Hyperbolic Skew T GARCH model(Aas and Haff (2006)). GHST1 is Generalized Hyperbolic Skew T GARCH model with Time Varying β. GHYT2 is Generalized Hyperbolic Skew T GARCH model with Time Varying ν. MIXNOR is a mixed normal GARCH model where the error is a mixture of two normal distributions. SAV is Symmetric Absolute Value CAViaR model(Engle and Manganelli (2004)) for quantiles. SSV is Symmetric Square Value CAViaR model. AS is Asymmetric Slop CAViaR model. AD is Adaptive CAViaR model.

	Estimation								
Simulation	NOR	STDT	SKEWT	GHST	MIXNOR	SAV	SSV	AS	AD
NOR	0.0072					0.0166	0.0385	0.0203	0.1546
	0.0187					0.0106	0.0317	0.0143	0.0441
STDT	0.0246	0.0304				0.0215	0.0550	0.0265	0.1694
	0.1336	0.1972				0.0153	0.0637	0.0193	0.0509
SKEWT	0.0339	0.0333	0.0120			0.0211	0.0565	0.0249	0.1734
	0.3629	0.3386	0.0451			0.0120	0.0529	0.0138	0.0489
GHST	0.0239	0.0240	0.0124	0.0091		0.1689	0.2576	0.1781	0.5635
	0.0762	0.0846	0.0432	0.0131		0.0596	0.1616	0.0636	0.1388
GHST1	0.1808	0.1735	0.0172	0.0158		0.0941	0.1737	0.1008	0.4402
	4.5877	3.3284	0.0976	0.0958		0.0816	0.2793	0.0890	0.1913
GHST2	0.0811	0.0342	0.0327	0.0243		0.0275	0.0457	0.0320	0.1912
	0.5213	0.0848	0.0323	0.0251		0.0523	0.2706	0.0688	0.2323
MIXNOR	0.0616	0.0591	0.0206	0.4184	0.0183	0.0240	0.0801	0.0310	0.1730
	0.1686	0.1662	0.0353	1.1890	0.0336	0.0089	0.0704	0.0170	0.0386

Table 2.7: Comparison of VaR using MSE

The entries are summary statistics(mean on the top and std on the bottom) of MSE to compare the simulation results of VaR. Simulations are conducted using 2500 observations across 1000 trials. The row names are simulation models, and the column names are estimation models. NOR is Gaussian GARCH model. STDT is TGARCH model. SKEWT is Skew T GARCH model(Hansen (1994)). GHST is Generalized Hyperbolic Skew T GARCH model(Aas and Haff (2006)). GHST1 is Generalized Hyperbolic Skew T GARCH model with Time Varying β. GHYT2 is Generalized Hyperbolic Skew T GARCH model with Time Varying ν. MIXNOR is a mixed normal GARCH model where the error is a mixture of two normal distributions. SAV is Symmetric Absolute Value CAViaR model(Engle and Manganelli (2004)) for quantiles. SSV is Symmetric Square Value CAViaR model. AS is Asymmetric Slop CAViaR model. AD is Adaptive CAViaR model.

	Estimation								
Simulation	NOR	STDT	SKEWT	GHST	MIXNOR	SAV	SSV	AS	AD
NOR	0.0051					0.0192	0.0569	0.0229	0.1527
	0.0236					0.0156	0.1081	0.0177	0.1026
STDT	0.0145	0.0094				0.0296	0.0962	0.0346	0.1783
	0.0308	0.0305				0.0488	0.5886	0.0439	0.2998
SKEWT	0.0125	0.0124	0.0097			0.0371	0.1071	0.0423	0.2131
	0.0341	0.0330	0.0357			0.0606	0.3099	0.0606	0.3450
GHST	0.0519	0.0254		0.0402		0.0675	0.1780	0.0782	0.2876
	0.0799	0.0483		0.0489		0.0941	0.4533	0.1218	0.3657
GHST1	0.0643	0.0289	0.0491	0.0344		0.0819	0.2667	0.0890	0.3730
	0.2766	0.3319	0.1946	0.1358		0.3900	2.1189	0.3898	1.4519
GHST2	0.0346	0.0088	0.0169	0.0041		0.0272	0.0586	0.0313	0.1467
	0.2294	0.0148	0.3161	0.0067		0.0729	0.3171	0.1088	0.6965
MIXNOR	0.0244	0.0129	0.0360	0.2085	0.0078	0.0412	0.1990	0.0467	0.2500
	0.0323	0.0211	0.0637	0.4420	0.0094	0.0528	0.4603	0.0495	0.2658

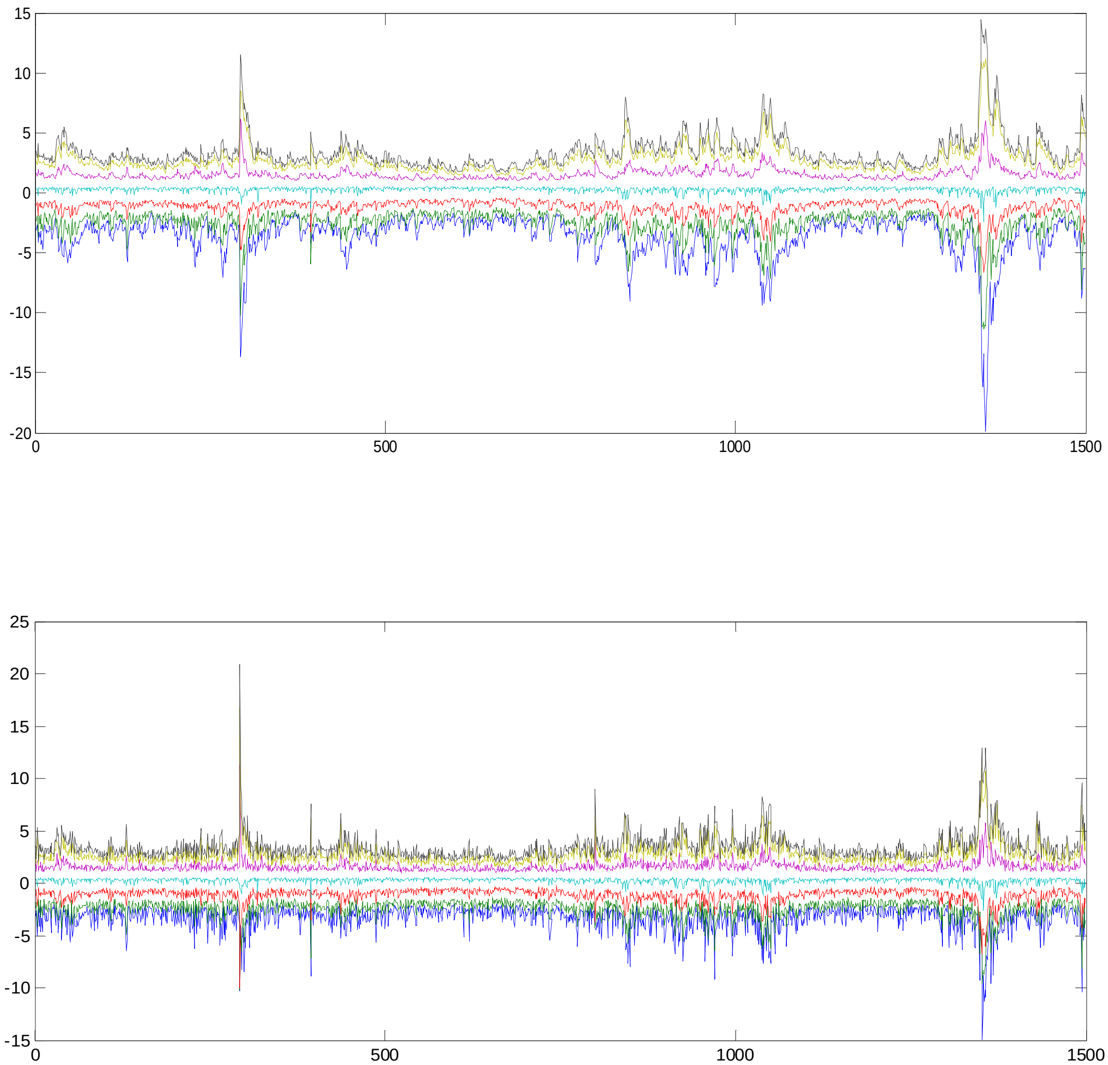

Figure 2.1: HYBRID quantile regression and MIDAS quantile regression: (a) the 5%, 10%, 25%, 50%, 75%, 90% and 95% quantiles for multiple horizon returns(horizon 5) using HYBRID quantile regression models with lag 5, (b) the 5%, 10%, 25%, 50%, 75%, 90% and 95% quantiles for multiple horizon returns(horizon 5) using MIDAS quantile regression models with lag 5.

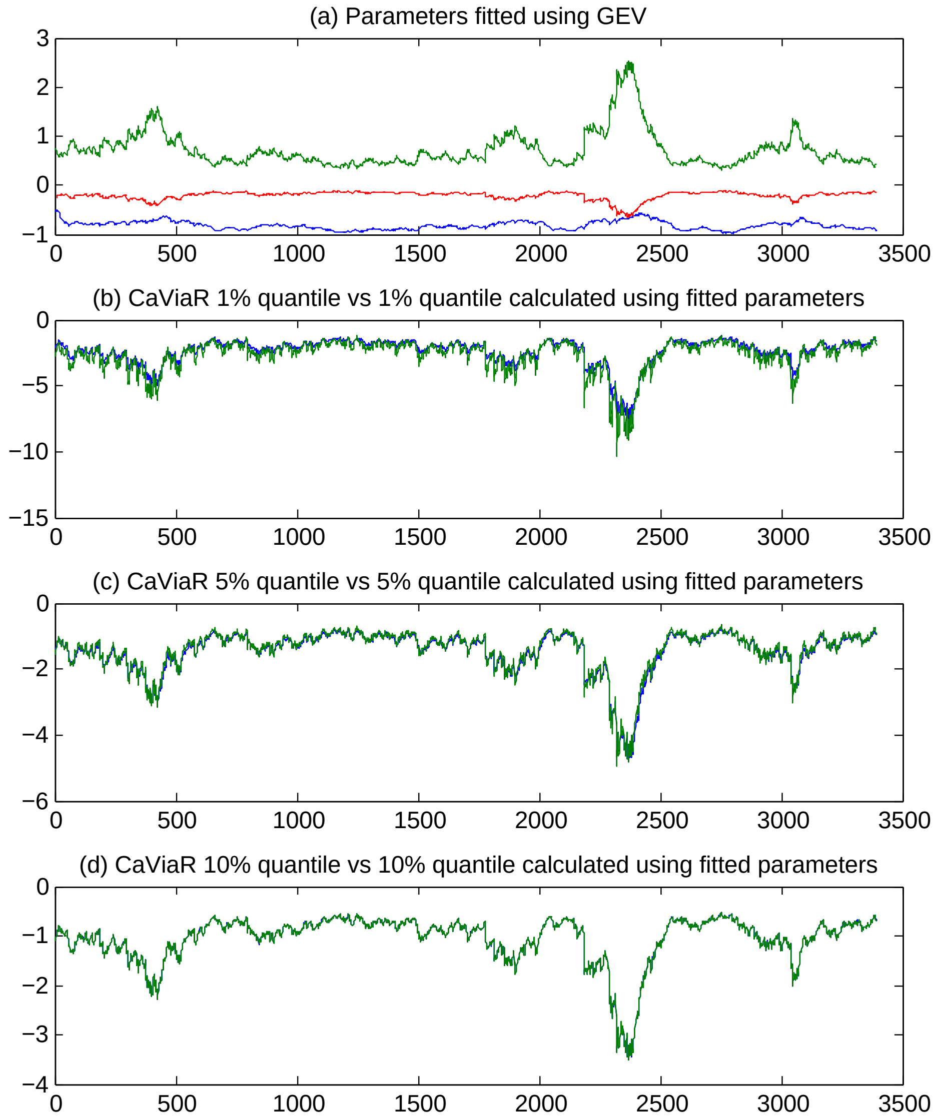

Figure 2.2: Comparison of quantiles by quantile distribution fits and CAViaR model: (a) the fitted parameters of generalized extreme value(GEV) distribution where the four quantiles (10% to 40% by 10%) used by quantile distribution fits are obtained by CAViaR SAV model using daily data, (b) CaViaR 1% quantile(Green) vs 1% quantile calculated using fitted parameters(Blue), (c) CaViaR 5% quantile vs 5% quantile calculated using fitted parameters, (d) CaViaR 10% quantile vs 10% quantile calculated using fitted parameters.

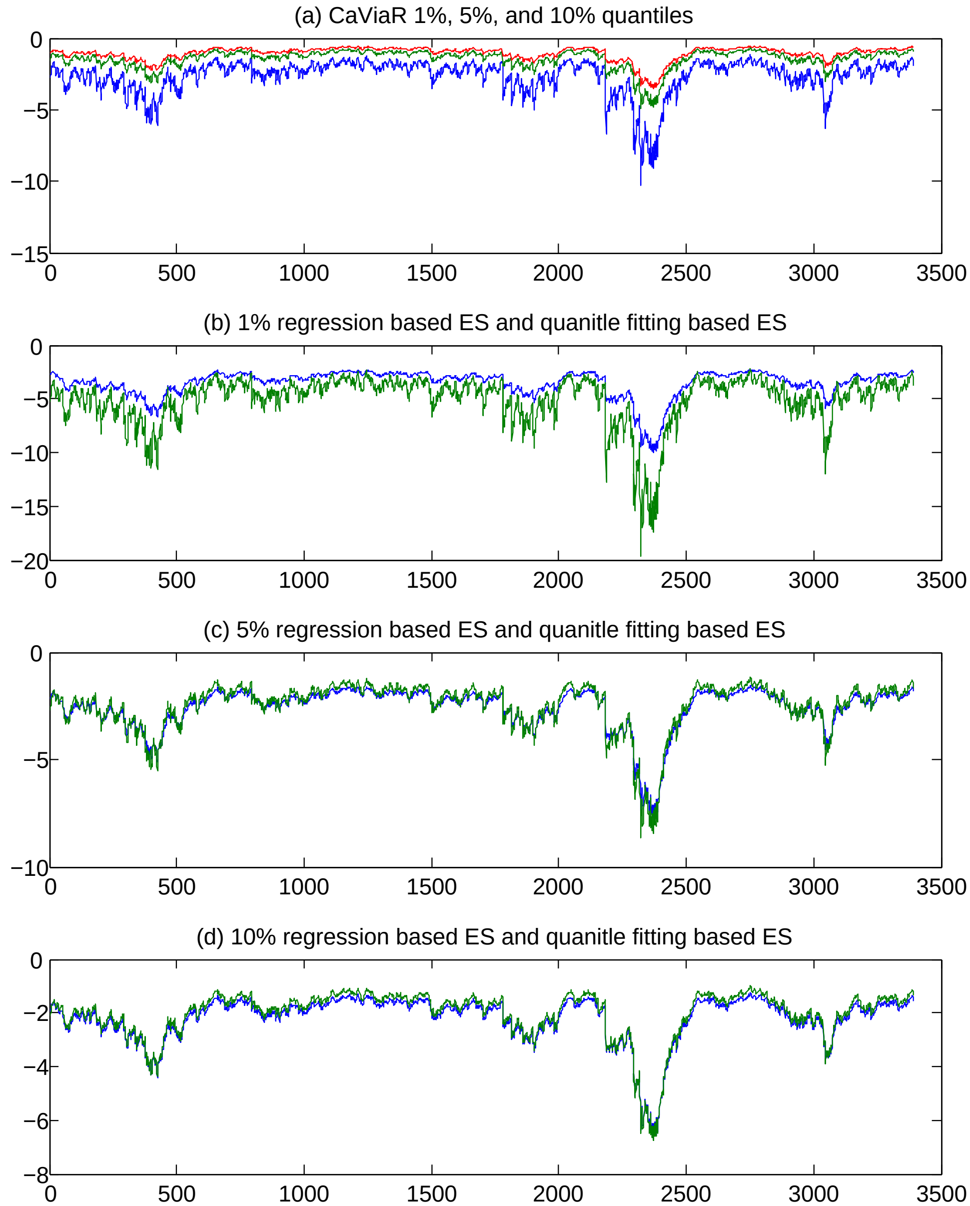

Figure 2.3: Comparison of Expected Shortfall(ES) by quantile distribution fits and regression based ES of CAViaR quantiles: (a) 1%, 5% and 10% CaViaR quantiles, (b) 1% regression based ES(Green) vs 1% quantile fitting based ES(Blue), (c) 5% regression based ES(Green) vs 5% quantile fitting based ES(Blue), (d) 10% regression based ES(Green) vs 10% quantile fitting based ES(Blue).

Chapter 3

Simulation Study of Long Run Skewness for Asset Pricing

3.1 Introduction

Bansal and Yaron (2004) have presented a consumption-based asset pricing model which includes a long-run predictable component, a time-varying consumption growth rates, time-varying volatility, and preference of Epstein and Zin (1989). Their model can explain some key features of dynamic asset pricing phenomena and address the asset market puzzles.

It has also been documented by empirical studies that the distribution of equity returns, either conditional or unconditional, can not be fully characterized by just mean and variance. Many previous studies have shown that the equity returns are negatively skewed(see e.g. Harvey and Siddique (2000)). Ghysels, Plazzi, and Valkanov (2010a) have also found a strong relationship between the conditional asymmetry and macroeconomic variables, which is different from the conditional volatility.

Inspired by these important findings, an intriguing question arises. Can we improve our understanding of equity returns and asset pricing by introducing higher moments into Bansal and Yaron (2004) type of model?

To better understand these questions, in this chapter, we are seeking to incorporate asymmetry in the Bansal and Yaron (2004) type of model and use simulation study to further investigate the long run skewness for an asymmetry consumption based asset pricing model that can generate larger equity returns due to asymmetry.

This chapter is structured as follows. Section 3.2.1 describes the asymetry consumption based asset pricing model. Section 3.2.2 provides the calibration of the model. Section 3.3 describes the simulation study using this model. Section 3.3.1 studies the Hansen Jagannathan Bound generated by this model. Section 3.3.2 provides distribution of equity returns for different parameter choices. Section 3.3.3 simulates the conditional moments of macro fundamentals and equity returns. In section 3.4, we conclude this chapter by sumerizing the findings.

3.2 Model Specification and Calibration

In this section, we first describe the threshold model of Colacito, Ghysels, Meng, and Ru (2012) in Section 3.2.1. Then the monthly calibration of the model is provided in Section 3.2.2.

Model Specification

Following Colacito, Ghysels, Meng, and Ru (2012), specify a representative consumer's preference at time t, U_t, as follows:

$$U_t = (1-\delta)\log C_t + \frac{\delta}{1-\gamma}\log E_t[\exp\{(1-\gamma)U_{t+1}\}] \tag{3.1}$$

Where γ is the degree of risk aversion, δ is the subjective discount factor, and C_t is the consumption at time t. This preference is the limiting case of Epstein and Zin (1989) when the intertemporal elasticity of substitution tends to be one. It is non time-additive while the constant relative risk aversion(CRRA) is time-additive. This preference has been used by several other papers, such as Colacito and Croce (2010), Kan (1995), Anderson (2005) and Lucas and Stokey (1984).

Let $\Delta c_t = \log(C_t) - \log(C_{t-1})$ denotes consumption growth. Following Colacito, Ghysels, Meng, and Ru (2012), we assume the consumption dynamic follows:

$$\Delta c_{t+1} = (\mu_c + \kappa_c) + \kappa_x x_t + \sigma_c \varepsilon_{c,t+1} \tag{3.2}$$

and the dividend growth $\Delta d_t = \log(D_t) - \log(D_{t-1})$ follows:

$$\Delta d_t = \lambda \Delta c_t \tag{3.3}$$

where $\lambda > 1$ is the leverage ratio for the claim on consumption and x_t is the long-run component of consumption growth which follows:

$$x_t = \rho_- x_{t-1} + \sigma_x \varepsilon_{x,t}, \ \forall x_{t-1} \leq 0 \tag{3.4}$$

$$x_t = \rho_+ x_{t-1} + \sigma_x \varepsilon_{x,t}, \ \forall x_{t-1} > 0 \tag{3.5}$$

Here, $\mu_c + \kappa_c$ is the average consumption growth, κ_x is the coefficient of x_t, σ_x is the volatility of shocks to x, σ_c is the standard deviation of the short-run shock to consumption, and ρ is autoregressive coefficient of long-run component x_t. For stationary, $\rho < 1$. The shocks $\varepsilon_{c,t}$ and $\varepsilon_{x,t}$ are i.i.d normal with mean zero and standard deviation 1. The model of Bansal and Yaron (2004) is a special case of the above model when $\rho = \rho_1 = \rho_+$, $\kappa_c = 0$, and $\kappa_x = 1$.

To solve the utility in equilibrium, we define the value function as follows:

$$V_t = U_t - \log C_t = \delta\theta \log E_t \exp \left\{ \frac{V_{t+1} + \Delta c_{t+1}}{\theta} \right\} \tag{3.6}$$

where $\theta = 1/(1-\gamma)$. Then the value function can be solved by iterating it on a grid of values of x_t.

For the preference given by 3.1, the stochastic discount factor, which is the intertemporal marginal rate of substitution, can be given as follows:

$$M_{t+1} = \frac{\partial U_t / \partial C_{t+1}}{\partial U_t / \partial C_t} \tag{3.7}$$

$$= \exp\left\{\log\delta - \Delta c_{t+1} + \frac{U_{t+1}}{\theta} - \log E_t \exp\left\{\frac{U_{t+1}}{\theta}\right\}\right\} \tag{3.8}$$

Let $m_t = \log M_t$ be the log consumption stochastic discount factor. The risk free rates can be written as:

$$r_t^f = -\log E_t \exp\{m_{t+1}\} \tag{3.9}$$

Define $v_{d,t} = P_t/D_t$ as price-dividend ratio(P/D ratio) and R_t^d as the returns to the dividend growth, which is levered consumption claim given by 3.3. The first order condition to price an asset implies that the return R_t^d satisfies Euler equation:

$$1 = E_t\left[M_{t+1}R_{d,t+1}\right] \tag{3.10}$$

Where the returns R_t^d is

$$R_{d,t+1} = \frac{P_{t+1} + D_{t+1}}{P_T} = \frac{1 + v_{d,t+1}}{v_{d,t}} \exp\{\Delta d_{t+1}\} \tag{3.11}$$

The log return is $r_{d,t+1} = \log R_{d,t+1}$. The dynamic of P/D ratio can be written as follows:

$$v_{d,t} = E_t\left[\exp\{m_{t+1}\}\left(1 + v_{d,t+1}\right)\exp\{\Delta d_{t+1}\}\right] \tag{3.12}$$

Calibration

Following Colacito, Ghysels, Meng, and Ru (2012), we calibrate the model at monthly frequency. The parameters choices are given by Table 3.1. The autoregressive coefficient ρ given in the table is for the benchmark case where $\rho = \rho_- = \rho_+$. Other choices of ρ_- and ρ_+ are listed in Table 3.2. The coefficient of risk aversion in Table 3.1 is set to 10 as a benchmark case. We study cases of γ from 7.5 to 20. The leverage is set to be 3 such that the dividend claim is more volatile than the consumption stream.

3.3 Simulation

After solving the value function, we simulate samples of length 100,000 with baseline parameter choices given by Table 3.1. Additional simulations are done for $\gamma \in \{7.5, 10, 12.5, 15, 17.5, 20\}$ with other parameters are same as Table 3.1 to study the relationship of $E[M]$ and $\sigma[M]$.

Section 3.3.1 studies the relationship between mean and variance of stochastic discount factor generated by this model. Section 3.3.2 provides distribution of equity returns for different parameter choices. Section 3.3.3 simulates the conditional moments of macro fundamentals and equity returns.

Hansen and Jagannathan Bound

Hansen and Jagannathan (1991) introduces Hansen and Jagannathan bounds which provide a criteria to validate whether a consumption based asset pricing model are feasible to compare asset pricing models. The Hansen and Jagannathan bounds are bound on the expectation of stochastic discount factor, standard deviation of the stochastic discount factor, and other moments of stochastic discount factor. Hansen and Jagannathan bound for a vector of returns, $\mathbf{R}$, is the hyperbola given by the following equation in $\{E[M], \sigma[M]\}$ space.

$$\sigma(M)^2 \geq (1 - E[M]E[\mathbf{R}])' \Sigma^{-1} (1 - E[M]E[\mathbf{R}]) \tag{3.13}$$

where Σ is the covariance matrix of $\mathbf{R}$.

Table 3.7 shows the results of pair of $E[M]$ and $\sigma[M]$.

Equity Returns

Table 3.2 shows the choice of parameters of ρ_- and ρ_+, and the means, volatilites, skewness, kurtosis, and first order autocorrelation of predictive component of consumption growth x_t, which follow the process of Equation 3.4 and 3.5. The choice of parameters of ρ_- and ρ_+ are chosen in 3.2 in order that the first order autocorrelation of consumption growth are the same across cases. We consider two choices of first order autocorrelation here, that is $\rho = 0.962$ and $\rho = 0.963$. To compare different cases, we need adjust κ_c and κ_x in order that the unconditional mean and volatility of consumption growth are the same across different cases(See Colacito, Ghysels, Meng, and Ru (2012)).

Table 3.3 shows the mean, variance, skewness, and kurtosis for both excess returns and risk free rates generated with parameters given by Table 3.1 and $\gamma = 15$. All numbers in the table are annualized. The first column is for baseline case with $\rho_- = \rho_+ = 0.962$. The simulated excess return has a mean of 2.391, and a slightly positive skewness. The larger the difference between ρ_- and ρ_+, the greater the expected excess return and negatively skewed. The risk free rates slightly decrease while the difference between ρ_- and ρ_+ increases. And the skewness of the risk free rates is always negative in the model from the simulations. The trends are the same for $\rho = 0.963$ cases.

Table 3.4 shows the same results with parameters given by Table 3.1 and $\gamma = 10$.

All numbers in the table are annualized. With $\gamma = 10$, the maximum expected excess return we can obtain from our selected parameters is 3.113. While for $\gamma = 15$, the maximum expected excess return we can obtain is 6.059, which is obtained when ρ_- and $\rho_=$ have the maximum difference.

From these we can conclude that the degree of asymmetry of autogressive coefficient of the long run component x_t plays an important role in the equity risk premia. That is, the degree of asymmetry of the predictive component of consumption growth largely determines the maximum Sharpe ratio that can be reached(Colacito, Ghysels, Meng, and Ru (2012)) and skewness can explains larger equity risk premia.

Table 3.5 shows the mean, variance, skewness, kurtosis for return, excess return, and risk free rates for parameters given by 3.1 and $\gamma = 15$ at multiple non-overlapping horizons from one month to one year. All numbers in the table are annualized. From this table, we can see that the variance is slightly reduced by aggregating with non-overlapping method, but the skewness is increased along the aggregating. We will show why this could be the case in Section 3.3.3 by evaluating the conditional moments of predictive component of consumption growth x_t and the conditional moments of excess returns. The variance of excess returns decreases while aggregating, and the skewness of excess returns increases. The skewness of risk free rates are larger than the skewness of excess returns, but the patterns are the same while aggregating.

Table 3.6 shows the same results for parameters given by 3.1 and $\gamma = 15$ at multiple overlapping horizons from one month to a year. All numbers in the table are annualized. All the patterns remains the same as aggregating using non-overlapping method.

Conditional Moments

Compared with Bansal and Yaron (2004), we introduce asymmetry in the predictive components of consumption growth rates x_t. Given our setting, the conditional skewness of $x_{t+1}|x_t$ should be zero, and for longer horizons, the distribution of conditional moments of $x_{t+n}|x_t$, where $n > 1$, are not clear. Hence, we simulate the $x_{t+n}, n \in \{1, ..., 12\}$ on a grid of x_t, which is equally spaced on the axis of x_t, for 10,000 times. Then, for each value on the grid of x_t, we calculate the expectation, variance, skewness, and kurtosis of x_{t+n}. These are the simulated conditional moments $x_{t+n}|x_t$. We are also interested in the conditional moments of excess returns. We simulate conditional moments of excess returns using the same method.

Figure 3.1 shows the conditional moments of $x_{t+n}|x_t$, where $n = 1, 3, 12$ for illustration. We can see that the conditional skewness of $x_{t+1}|x_t$ is zero and conditional variance is constant as expected. The conditional skewness of $x_{t+n}|x_t$ is increasing while the number of horizons n increases, especially when x_t is near zero. This is the case since the asymmetry we introduce in the model is indeed a threshold model while the threshold is at zero.

Figure 3.2 shows the conditional moments of $r_{t+n}|x_t$, where $n = 1, 3, 12$ for illustration. All the numbers in the figure are annualized. The same pattern holds as the conditional moments of $x_{t+n}|x_t$. The conditional excess returns attain the maximum at $x_t = 0$.

3.4 Conclusion

By introducing asymmetry in the autoregressive coefficient of the long run component x_t (predictive component of consumption growth rates), therefore asymmetry in the predictive component of consumption growth rate, we propose an asymmetry version of Bansal and Yaron (2004). We study the relationship between the expected stochastic discount factor and variance of the stochastic discount factor. As shown by Colacito, Ghysels, Meng, and Ru (2012), the Hansen and Jagannathan bound can be attained and larger Sharp ratio can also be achieved.

By increasing the asymmetry in the predictive component of consumption growth rates, larger expected excess returns can be obtained. And the skewness of both excess return and risk free rates increase as the asymmetry in the autoregressive coefficient of the long run component increases. We also study the distribution of the excess return and risk free rates over longer horizon by overlapping and non-overlapping methods. The results show that the variance slightly decreases while the horizon increases and the skewness increases for both excess returns and risk free rates using both overlapping and non-overlapping aggregating methods.

By introducing asymmetry in the predictive component of consumption growth rates x_t, the conditional moments of x_t becomes time-varying at multiple horizons when aggregating without overlapping. The conditional distribution of $x_{t+n}|x_t$ become time-varying, and more negatively skewed. The conditional moments for excess returns also become more negatively skewed when increasing horizon.

Given the inspiring findings in this chapter, one can expect to explain larger excess returns using the consumption based asset pricing by introducing conditional asymmetry in the long run component of consumption growth rates. Therefore, conditional asymmetry/ conditional skewness may offer a promising approach to address equity premium puzzle and could significantly improve our understanding on the risk management and portfolio selection in the future.

3.5 Tables and Figures

The following are Tables and Figures of this chapter.

Table 3.1: Monthly Calibration

γ	Risk aversion	10 or 15
δ	Subjective discount factor	0.9989875
μ_c	Average consumption growth	0.001
ρ		0.962 or 0.963
κ_c		0
κ_x		1
σ_c	Standard deviation of the short-run shock to consumption	0.0068
σ_x	Volatility of shock to x	$0.05\sigma_c$
λ	Leverage	3

Table 3.2: Distribution of Predictive Components for Monthly Calibration

ρ_-	ρ_+	$E[x]$	$\sigma[x]$	$skew[x]$	$kurt[x]$	$\rho[x_t, x_{t-1}]$
0.962	0.962	0.000		0.000	3.000	0.962
0.972	0.945	-0.978	3.674	-0.254	3.047	0.962
0.980	0.868	-2.470	3.654	-0.605	3.288	0.962
0.981	0.841	-2.716	3.662	-0.653	3.337	0.962
0.963	0.963	0.000		0.000	3.000	0.963
0.976	0.930	-1.531	3.704	-0.387	3.113	0.963
0.978	0.915	-1.891	3.713	-0.470	3.167	0.963
0.979	0.899	-2.138	3.695	-0.528	3.212	0.963
0.980	0.899	-2.351	3.710	-0.574	3.252	0.963
0.981	0.874	-2.531	3.744	-0.606	3.287	0.963
0.981	0.858	-2.632	3.699	-0.634	3.316	0.963
0.982	0.834	-2.860	3.727	-0.673	3.361	0.963

Table 3.3: Equity return for $\gamma = 15$

ρ_-	0.962	0.972	0.980	0.981	0.963	0.976	0.978	0.979	0.980	0.981	0.982
ρ_+	0.962	0.945	0.868	0.841	0.963	0.930	0.915	0.899	0.899	0.858	0.834
$E\left[r_d - r_f\right]$	2.391		5.162	5.551	2.510	4.060	4.572	4.883	5.236	5.649	6.059
$\sigma\left[r_d - r_f\right]$	9.212		11.599	11.841	9.304	10.731	11.131	11.368	11.598	11.859	12.069
$skew\left[r_d - r_f\right]$	0.002		-0.036	-0.035	0.002	-0.029	-0.032	-0.033	-0.033	-0.034	-0.032
$kurt\left[r_d - r_f\right]$	3.009		3.186	3.207	3.010	3.091	3.123	3.148	3.165	3.191	3.204
$E\left[r_f\right]$	2.382		2.365	2.365	2.382	2.372	2.369	2.367	2.366	2.362	2.361
$\sigma\left[r_f\right]$	0.435		0.441	0.441	0.440	0.445	0.445	0.445	0.446	0.446	0.0446
$skew\left[r_f\right]$	-0.036		-0.620	-0.667	-0.038	-0.410	-0.491	-0.547	-0.589	-0.647	-0.687
$kurt\left[r_f\right]$	2.948		3.276	3.327	2.947	3.100	3.159	3.206	3.245	3.304	3.348

Table 3.4: Equity return for $\gamma = 10$

ρ_-	0.962	0.972	0.980	0.981	0.963	0.976	0.978	0.979	0.980	0.981	0.982
ρ_+	0.962	0.945	0.868	0.841	0.963	0.930	0.915	0.899	0.899	0.858	0.834
$E\left[r_d - r_f\right]$	1.505	1.859	2.641	2.819	1.582	2.217	2.436	2.559	2.724	2.905	3.113
$\sigma\left[r_d - r_f\right]$	9.290	9.897	11.163	11.422	9.390	10.454	10.797	10.997	11.236	11.503	11.785
$skew\left[r_d - r_f\right]$	0.002	-0.027	-0.058	-0.060	0.002	-0.041	-0.048	-0.053	-0.055	-0.059	-0.059
$kurt\left[r_d - r_f\right]$	3.010	3.060	3.282	3.324	3.010	3.127	3.182	3.224	3.259	3.309	3.343
$E\left[r_f\right]$	2.382	2.376	2.365	2.365	2.382	2.372	2.369	2.367	2.366	2.362	2.361
$\sigma\left[r_f\right]$	0.435	0.438	0.441	0.441	0.440	0.445	0.445	0.445	0.446	0.446	0.446
$skew\left[r_f\right]$	-0.036	-0.281	-0.620	-0.667	-0.038	-0.410	-0.491	-0.547	-0.589	-0.647	-0.687
$kurt\left[r_f\right]$	2.948	3.027	3.276	3.327	2.947	3.100	3.159	3.206	3.245	3.304	3.348

Table 3.5: Multihorizon equity return for $\gamma = 15$

	Non-overlapping Horizon											
	1	2	3	4	5	6	7	8	9	10	11	12
$E[r_d]$	8.419	8.419	8.419	8.419	8.419	8.420	8.420	8.419	8.419	8.419	8.420	8.420
$\sigma[r_d]$	12.149	12.135	12.163	12.089	12.024	12.024	11.943	12.018	11.990	11.925	11.995	11.978
$skew[r_d]$	-0.052	-0.063	-0.079	-0.077	-0.094	-0.096	-0.140	-0.146	-0.168	-0.120	-0.174	-0.198
$kurt[r_d]$	3.199	3.136	3.156	3.183	3.094	3.077	3.023	3.126	3.109	3.085	3.096	3.088
$E[r_d - r_f]$	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.060	6.059
$\sigma[r_d - r_f]$	12.069	12.022	12.018	11.912	11.816	11.785	11.674	11.718	11.660	11.570	11.615	11.564
$skew[r_d - r_f]$	-0.032	-0.035	-0.045	-0.035	-0.049	-0.048	-0.086	-0.086	-0.108	-0.051	-0.100	-0.132
$kurt[r_d - r_f]$	3.204	3.142	3.167	3.197	3.114	3.098	3.042	3.146	3.127	3.130	3.126	3.098
$E[r_f]$	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361
$\sigma[r_f]$	0.446	0.626	0.761	0.874	0.971	1.057	1.136	1.208	1.274	1.335	1.393	1.447
$skew[r_f]$	-0.687	-0.706	-0.720	-0.732	-0.744	-0.756	-0.764	-0.776	-0.786	-0.795	-0.802	-0.812
$kurt[r_f]$	3.348	3.359	3.369	3.380	3.389	3.401	3.402	3.424	3.425	3.449	3.460	3.470

Table 3.6: Multihorizon equity return for $\gamma = 15$

	Overlapping Horizon											
	1	2	3	4	5	6	7	8	9	10	11	12
$E[r_d]$	8.419	8.419	8.419	8.419	8.420	8.420	8.420	8.420	8.420	8.419	8.419	8.419
$\sigma[r_d]$	12.149	12.135	12.129	12.109	12.072	12.044	12.017	11.993	11.973	11.954	11.942	11.929
$skew[r_d]$	-0.052	-0.065	-0.071	-0.087	-0.106	-0.120	-0.133	-0.144	-0.150	-0.158	-0.167	-0.177
$kurt[r_d]$	3.199	3.122	3.119	3.109	3.112	3.110	3.107	3.102	3.102	3.094	3.087	3.082
$E[r_d - r_f]$	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059	6.059
$\sigma[r_d - r_f]$	12.069	12.022	11.984	11.932	11.864	11.806	11.749	11.695	11.646	11.598	11.559	11.519
$skew[r_d - r_f]$	-0.032	-0.037	-0.036	-0.047	-0.062	-0.071	-0.079	-0.085	-0.087	-0.091	-0.097	-0.102
$kurt[r_d - r_f]$	3.204	3.130	3.133	3.124	3.130	3.132	3.132	3.130	3.132	3.126	3.119	3.116
$E[r_f]$	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361	2.361
$\sigma[r_f]$	0.446	0.626	0.761	0.874	0.971	1.057	1.136	1.208	1.274	1.335	1.393	1.447
$skew[r_f]$	-0.687	-0.706	-0.720	-0.732	-0.744	-0.756	-0.764	-0.776	-0.786	-0.795	-0.802	-0.812
$kurt[r_f]$	3.348	3.359	3.369	3.380	3.389	3.401	3.402	3.424	3.425	3.449	3.460	3.470

Table 3.7: Stochastic discount factor

	γ	7.5		10		12.5		15		17.5	
ρ_-	ρ_+	$E[M]$	$\sigma[M]$	$E[M]$	$\sigma[M]$	$E[M]$	$\sigma[M]$	$E[M]$	$\sigma[M]$	$E[M]$	$\sigma[M]$
0.962	0.962			0.998	0.104	0.999	0.132	0.999	0.160	0.999	0.188
0.972	0.945	0.998	0.079	0.999	0.110					0.999	0.207
0.980	0.868	0.998	0.085	0.999	0.122	0.999	0.162	0.999	0.205	0.999	0.250
0.981	0.841	0.998	0.087	0.999	0.124	0.999	0.168	0.999	0.211	0.999	0.262
0.963	0.963			0.996	0.184	0.996	0.234	0.997	0.285	0.997	0.337
0.976	0.930	0.995	0.145	0.996	0.203	0.996	0.266	0.997	0.334	0.998	0.406
0.978	0.915	0.995	0.148	0.996	0.210	0.997	0.277	0.997	0.351	0.998	0.429
0.979	0.899	0.996	0.150	0.996	0.213	0.997	0.284	0.997	0.361	0.998	0.444
0.980	0.899	0.996	0.152	0.996	0.218	0.997	0.292	0.997	0.373	0.998	0.461
0.981	0.858	0.996	0.154	0.996	0.223	0.997	0.301	0.998	0.387	0.998	0.481
0.982	0.834			0.996	0.229	0.997	0.311	0.998	0.402	0.998	0.502

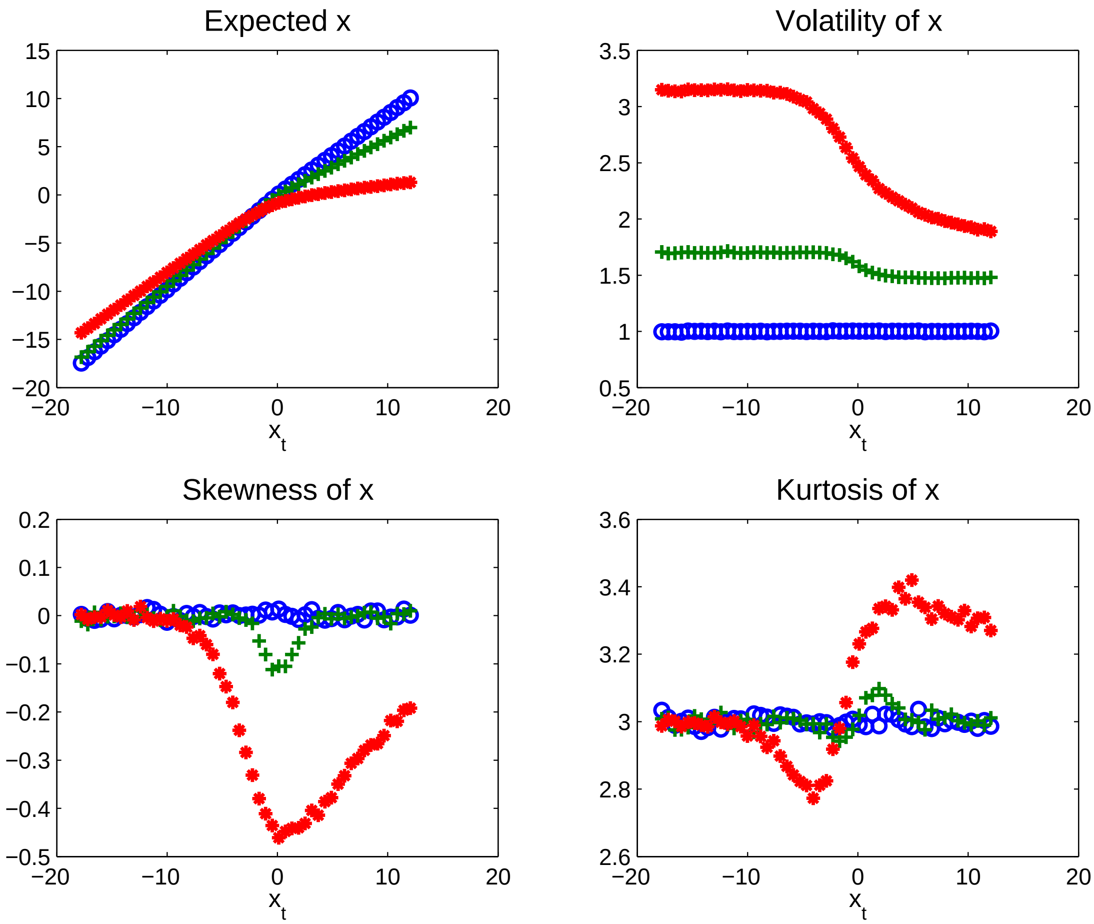

Figure 3.1: Conditional Moments of x_t for multiple horizons: moments of $x_{t+1}|x_t$ in blue, $x_{t+3}|x_t$ in green and $x_{t+12}|x_t$ in red

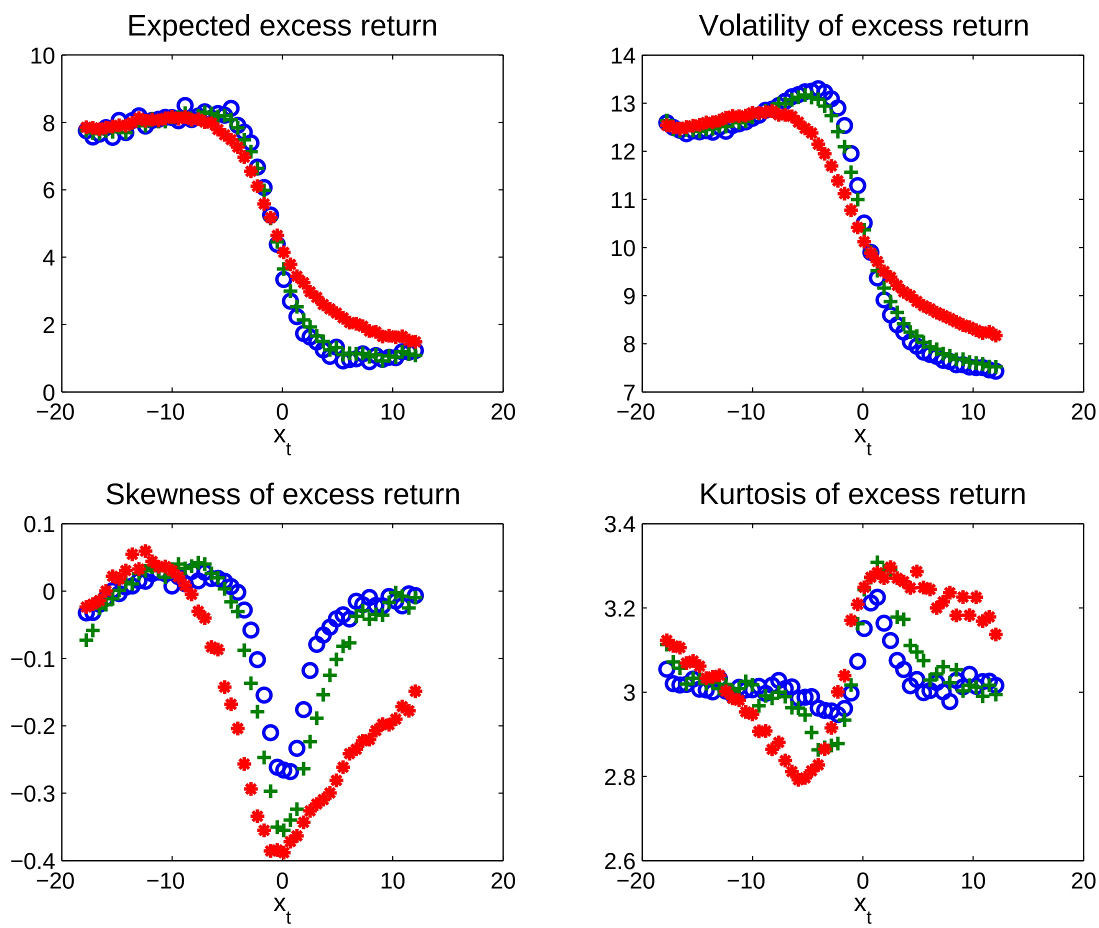

Figure 3.2: Conditional Moments of excess return for multiple horizons: moments of $r_{e,t+1}|x_t$ in blue, $r_{e,t+3}|x_t$ in green and $r_{e,t+12}|x_t$ in red

Bibliography

Aas, K., and I. H. Haff, 2006, The generalized hyperbolic skew student's t-distribution, *Journal of Financial Econometrics* 4, 275–309.

Anderson, E., 2005, The Dynamics of Risk-sensitive Allocations, *Journal of Economic Theory* 125, 93–150.

Artzner, P., F. Delbaen, J. Eber, and D. Heath, 1997, Thinking Coherently, *Risk* 10(11), 68–71.

Bansal, R., and A. Yaron, 2004, Risk for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 4, 1481–1509.

Bekaert, G., and Eric Engstrom, 2010, Aseet Return Dynamics under Bad Environment-Good Environment Fundamentals, Working Paper.

Bowley, A., 1920, *Elements of Statistics* (New York: Charles Scribner's Sons).

Burnside, C., 1998, Solving Asset Picing Models with Gaussian Shocks, *Journal of Economic Dynamics and Control* 22, 329–340.

Campbell, J.Y., and J.H. Cochrane, 1999, By force of habit: A consuption-based explanation of aggregate stock market behavior, *Journal of Financial Economics* 107, 205–251.

Chen, X., E. Ghysels, and F. Wang, 2010, HYBRID-GARCH: A Generic Class of Models for Volatility Predictions using Mixed Frequency Data, Working Paper, UNC.

Colacito, R., and M.M. Croce, 2010, International Asset Pricing with Risk Sensitive Rare Events, Working Paper, UNC.

Colacito, R., E. Ghysels, J. Meng, and H. Ru, 2012, Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory, Working Paper.

Engle, Robert, and Simone Manganelli, 2004, Caviar: Conditional autoregressive value at risk by regression quantiles, *Journal of Business and Economic Statistics* 22, 367–381.

Engle, Robert, and Abhishek Mistry, 2007, Priced risk and asymmetric volatility in the cross-section of skewness, *Working Paper, NYU Stern School of Business.*

Epstein, L.G., and S.E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Escanciano, J. C., 2009, QMLE Estimation of Semi-strong GARCH Models, *Journal of Economic Theory* 25, 561–570.

Francq, C., and J.M. Zakoian, 2004, Maximum likelihood estimation of pure garch and arma-garch process, *Bernoulli* 10, 605–637.

Ghysels, E., A. Plazzi, and R. Valkanov, 2010a, Conditional Skewness of Stock Market Returns in Developed and Emerging Markets and its Economic Fundamentals, Working Paper.

——— , 2010b, On the Term Structure of Conditional Skewness in Stock Returns, Work in progress.

——— , 2011, Conditional skewness of stock market returns in developed and emerging markets and its economic fundamentals, Paper available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1761446`.

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2006, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* 131(1-2), 59–95.

Gouriéroux, C., and J. Jasiak, 2008, Dynamic quantile models, *Journal of Econometrics* 147(1), 198–205.

Groeneveld, R., and G. Meeden, 1984, Measuring skewness and kurtosis, *The Statistician* 33, 391–399.

Haas, M., S. Mittnik, and M. S. Paolella, 2004, Mixed Normal Conditional Heteroskedasticity, *Journal of Financial Econometrics* 2(2), 211–250.

Hall, P., and Q. Yao, 2003, Inference in ARCH and GARCH models with heavy–tailed errors, *Econometrica* 71(1), 285–317.

Hansen, B. E., 1994, Autoregressive Conditional Density Estimation, *International Economic Review* 35, 705–730.

Hansen, L.P., and R. Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy* 99(2), 225–262.

Harvey, C.R., and A. Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* LV, 1263–1295.

Kan, R., 1995, Structure of Pareto Optima When Agents Have Stochastic Recursive Preferences, *Journal of Economic Theory* 66, 626–631.

Kim, Tae-Hwan, and Halbert White, 2004, On more robust estimation of skewness and kurtosis, *Finance Research Letters* 1, 56–70.

Koenker, R., and G. Bassett, 1978, Regression quantiles, *Econometrica* 46, 33–50.

Koenker, R., and Z. Xiao, 2006, Quantile autoregression, *Journal of the American Statistical Association* 101(475), 980–990.

Komunjer, I., 2004, Quasi-maximum Likelihood Estimation for Conditional Quantiles, *Journal of Econometrics* 128, 137–164.

Lucas, R.E., and N.L. Stokey, 1984, Optimal growth with many consumers, *Journal of Economic Theory* 32, 139–171.

Manganelli, S., and R. F. Engle, 2001, Value at Risk Models in Fiance, Working Paper.

Patton, A., 2011, Data-based ranking of realised volatility estimators, *Journal of Econometrics* 161, 284–303.

Sakata, S., and H. White, 1998, High breakdown point conditional dispersion estimation with application to S & P 500 daily returns volatility, *Econometrica* 66(3), 529–567.

Tsionas, E. G., 2003, Exact Solution of Asset Pricing Models with Arbitrary Shock Distributions, *Journal of Economic Dynamics and Control* 27, 843–851.

Weiss, A.A., 1991, Estimating nonlinear dynamic models using least absolute error estimation, *Econometric Theory* 7(01), 46–68.

White, H., 1996, *Estimation, inference and specification analysis* (Cambridge University Press).

White, Halbert, Tae-Hwan Kim, and Simone Manganelli, 2008, Modeling autoregressive conditional skewness and kurtosis with multi-quantile caviar, in R. F. Engle, and H. White, ed.: *A Festschrift in Honor of Robert F Engle*. Oxford University Press.

——— , 2010, VAR for VaR: Measuring Systemic Risk Using Multivariate Regression Quantiles, Discussion Paper, ECB and UCSD.

Wu, X., 2006, Inference and Density Estimation with Interval Statistics, Discussion Paper, Department of Agricultural Economics, Texas A & M University.

——— , and J.M. Perloff, 2005, China's income distribution, 1985-2001, *Review of Economics and Statistics* 87, 763–775.

——— , 2007, GMM estimation of a maximum entropy distribution with interval data, *Journal of econometrics* 138, 532–546.