On the Robustness of Beta Risk In The Cross-Sectional
Regressions Of Stock Return

Chuanliang Jiang

Dec 7th, 2011

Abstract

Financial data are typically characterized by skewness, excess kurtosis or in general deviation from normality. Traditional OLS estimation of beta in CAPM model has been criticized because of non-robustness of estimation. In this paper, we propose a novel quantile-based approach to estimate the conditional beta risk more efficiently without assuming a parametric structure on distribution. Multiple quantiles estimates are combined in weighting schemes to utilize information across the whole distribution. Monte Carlo simulation demonstrates significant efficiency improvement for beta estimation from our proposed approach when tail of distribution is thicker than normal distribution. Applying our quantile-based method in the financial market to estimate firm specific beta risk in our sample data, a pronounced difference was found between OLS and quantile combination estimator. Our robust approach also applied to Fama MacBeth Capital Asset Pricing Test. The goodness of fit for the Fama MecBeth second pass cross sectional regression increase slightly as we use quantile combination approach instead of OLS to estimate beta risk, which implies a lower measurement error of beta leading to the improvement in the second pass cross section regression. Finally, we studied the robust measurement of factor risk Premium and found that the size factor effect disappear after introducing our robust approach to estimate risk premium

JEL Classifications: C21 G12
1. Introduction

The Beta of a security represents assets’ systematic risk which can’t be diversified away by constructing a portfolio of many risky asset. The most common asset pricing model (CAPM) predict that only beta risk is priced in equilibrium, and a stock is riskier the more closely its prices co-move with market prices as a whole. Precise estimation of beta risk play a crucial role in equity valuation, risk management and asset allocation. Slight variation in the beta estimates would lead to diverging company share valuation. Given the imprecision estimation of beta risk, portfolio manager would fail to calculate the cost of equity and minimize risk to reward ratios.

Since beta risk is latent variable which is unobservable to investors, various estimation strategies had been proposed attempting to obtain more accurate estimator of beta as proxy for its true value. A traditional way to increase precision of beta is to group stocks into portfolios and estimate the beta for each portfolio in time series regression. However, some recent studies, like Ang, Liu and Schwarz (2010), Cosemans, Frehen, Schotman and Bauer (2011) advocate the use of individual firms as base assets. They argued that constructing portfolio would conceal important information contained in individual stock betas, reducing the cross sectional variation in beta estimators. Especially, the extreme outliers influences would be disguised in portfolio construction. An alternative way to improve the precision of beta estimation is to introduce more sophisticate estimation strategies attempting to model time variability behavior of beta risk. Adrian and Franzoni (2005), Petkova and Zhang (2005), Lewellen and Nagel (2006), Ang and Chen (2007), among others all documented that constant beta estimates fail to capture investor characteristics and may lead to inaccurate estimates of the true underlying beta. Although these studies have achieved great success in modeling the true behavior of beta risk, an important drawback of these framework is that they all assume the assumption of normal distribution without testing its reliability, thus ignoring the limitations of model in the case of non-normal distribution.

Many empirical studies, however, have found that the return distribution of security

\footnote{Since the literature in this part is too large, it is impossible to highlight all of related papers here.}
are significantly non-normal with thick tail (leptokurtosis) and skewness. Extreme events occur much more frequently than that predicted by the normal distribution. Mandlebrot (1963) and Fama (1965) firstly recognized that stock return residuals have thick tails. Since then, Clark (1973), Bollerslev (1987), Hensen (1991), among others, found as well that the departure of normal distribution is not rare at all in the financial data. More recently, Harvey and Siddique (1999) found skewness and fat tails in tests of various stock indices and assets. Recent research suggests that higher moments such as skewness and kurtosis are priced in asset pricing test. Besides, Hedge fund literatures documented a significant deviation from normality for Hedge fund data which can be attributed to its highly dynamic complex trading strategies (Hamidi, Maillet and Merlin (2011)). In spite of this, the normality assumption is still the working assumption in the mainstream finance. Kan and Zhou (2006) claimed that: "the reason for the wide use of the normality assumption is not because it model financial data well, but due to its tractability that allows interesting economic questions to be answered ...". And they advocated the use of symmetric student t distribution to fit financial data and found that the implication of asset pricing test dramatically altered.

It is well established that in the small sample, ordinary least square estimation are sensitive to outlier observation. Estimating CAPM model using ordinary lease square (OLS) would result in dramatic loss of efficiency and large dispersion of measurement error when the data follow fat tailed non-normal distribution. It is therefore essential to investigate a robust estimation of beta risk under non-normality of security return.

Recognizing the weakness of normal distribution, Various studies have proposed alternative distributions attempting to fit financial data well. Clark (1973), McCulloch (1985), Bollerslev (1987), Nelson (1991), Hansen (1994), Mitnik et al. (1999), Kan and Zhou (2006), (McDonald Michelfelder and Theodossiou (2009)), among others, suggests t distribution, skew t-distribution, general error distribution (GED), \( \alpha \)-stable levy distribution or some other flexible probability distributions. However, these distribution are either, by no means parsimonious or fail to measure the true distribution function form of financial data, which is probably unknown to the econometricians. Allowing for a parametric structure on the unknown distribution form would lead to misleading inferences if this parametric representation is misspecified.
In this paper, we proposed a novel and sophisticated quantile-based approach to estimate a robust beta risk in the presence of fat tail distribution. Multiple quantiles estimates are combined in several weighting schemes to utilize information across the whole distribution. Our Monte Carlo simulation below demonstrated that combining multiple quantile estimates can substantially improve the efficiency of estimation (lower measurement error) for beta risk when tail of distribution is thicker than that of normal distribution. This efficiency gain still prevails even in the asymmetric distribution with positive (negative) skewness. Applying our robust approach to estimate beta risk for individual firms in our sample data, we found a significant difference between quantile-based estimator and OLS estimator. More generally, OLS estimation of beta based on our sample monthly (daily) data seems to overestimate (underestimate) the true value of beta under fat tail distribution of data. Besides, we demonstrate the economic significance of our robust estimator of beta risk through Fama MecBeth two pass asset pricing test. With the departure of normality distribution, our quantile-based approach is expected to yield a more precise estimation of beta, leading to the improvement in the second-pass cross-sectional regression.

Meanwhile, we suspect that the motivation to combine multiple quantiles not only come from its statistics robustness to outlier observation under fat tail distribution, but also due to its necessity to aggregate the information of different causal effect or sensitivity between response and factors over various quantiles. Baur and Schulze (2010) documented the beta of individual stocks varies across the entire return distribution. They employed a quantile regression framework to estimate the regime-dependent systematic beta risks on individual stocks and explained time varying behavior of beta with investor learning for daily data and prospect theory for monthly data. Their empirical studies imply that the risk of individual stock can be underestimated or overestimated significantly if the analysis is confined to conditional means or a specific quantile (regime). On the contrary, our quantile combination approach facilitates the utilization of distribution information across different quantiles.

It can’t be denied that our paper is not the first study regarding quantile combination to estimate beta. One of papers most closely related to ours is from Chan

\[ \beta(\tau) = \beta + F^{-1}_\tau(\tau) \]

\(^2\)It is equivalent to combine beta risk in the conditional heteroskedasticity model where beta has quantile heteroskedasticity.
& Lakonishok(1992) who found efficiency gain for the estimation of CAPM model from simple trimmed regression quantiles (TRQ). However, their studies was purely based on simulation exercise and very premitive in the sense that they were constructed under the "ideal" situation where only two extreme distribution forms are considered: normal distribution and symmetric and heavily tailed t distribution with $df = 3$. Our studies, however, are more complete and sophisticated. We study the robustness of quantile combination under different shape of distribution and found that skewness or conditional heteroskedasticity may "dilute" the efficient gain of parameter estimation from quantile combination if no corrections are implemented. Moreover, the efficiency improvement from quantile combination estimation is more readily available for the slope coefficient $\beta$ than for the intercept $\alpha$, which is more likely to be influenced by the empirical quantile of residuals.

We proposed an alternative and more sophisticate density weighted quantile combination approach in which the weight assigned to each quantile are really data-driven and determined by the quantile heteroskedasticity component. The performance of various quantile combination approach are compared and evaluated both in simulation exercises and empirical studies. More recent paper from Benjamin, Bertrand and Paul(2011) introduced a time varying multi-quantile framework to study the style analysis for Hedge Fund industry. Their studies, however, is still not satisfactory due to its lacks of detailed analysis in quantile combination.

The remaining of paper proceeds as follows. In section 2, we briefly introduce the estimation of beta risk in financial literature. In section 3, we establish the model of quantile combination approach and develop the method to estimate weight average to combine quantile. In section 4, we construct Monte Carlo simulation to evaluate the performance of quantile combination under different shape of distribution. In section 5, we applies the quantile combination method in empirical studies to check if quantile combination produce robust estimator of beta risk in real data. And section 6 concludes that paper...
2. Capital Asset Pricing Model

The capital asset pricing model (CAPM) states that the expected return of asset $i$ is directly related to the beta risk of overall stock market.

$$E_t(r_{i,t}) = \lambda_{t+1} \beta_{i,t}$$  \hspace{1cm} (1)

where $r_{i,t} = R_{i,t+1} - R_{f,t+1}$ is excess stock return. $\lambda_{t+1}$ is risk premium for the exposure to market risk. Since systematic risk $\beta$ is latent variable and unobservable, many literature had propose various methodology to estimate beta risk. Fama and MacBeth (1973) had suggested two-pass procedure in which beta estimates are obtained firstly in the first stage of time series regression for each underlying asset

$$r_{i,t} = \alpha_{i,t} + \beta_{i,t} r_{M,t} + \epsilon_{i,t}$$

Because of the unknown value of $\beta_t$, a predictor of beta $\hat{\beta}_{t-1}$ will be used as a proxy for the unobservable true value of $\beta$, which may induce the predictive error in the second stage of cross sectional regression. More generally, the beta risk is measured with an error

$$\hat{\beta}_{t-1} = \beta_t + \xi_{t-1} \quad \xi_{t-1} \sim N(0, \sigma^2_{\xi,t-1})$$

Shanken (1992) argued that the measurement error of beta risk could create serious bias in the small sample of cross sectional asset pricing test. The standard method of dealing with this problem is to aggregate individual firms into portfolio. The information content inherent in the individual returns data, however, will be lost significantly. Ang & Liu (2010) examined the efficiency of using individual stock or portfolios as based assets and found that the smaller standard error of beta estimates from portfolio construction do not lead to smaller standard errors of cross sectional coefficient estimates. In this paper, we also use individual securities as base asset to estimate beta since portfolio grouping can distort the impact of extreme outliers in the fat tailed distribution.

With knowledge of the shape of distribution, parameters of linear model can be estimated most efficiently by maximal likelihood estimation (MLE).

$$[\alpha, \beta|\theta_i] = \arg\max_{\alpha,\beta} \left( \sum_{t=1}^{T} \ln f_i(r_{i,t} - \alpha_{i,t} - \beta_{i,t} r_{M,t}|\theta_i) \right)$$  \hspace{1cm} (2)
However, MLE was either infeasible due to the unknown parametric form of model, or intractable due to computation difficulty. Most of asset pricing literatures, thus, assume stock return follow normal distribution, even though it is widely accepted that stock and other asset return pdfs have fat tail due to extreme outlier observation. With the distribution of non-normality, OLS estimation may result in inefficient estimator of beta risk which tends to have greater estimation error.

Following the Fama and MacBeth two-pass Methodology (1973), we estimate quantiles of the entire conditional distribution for beta risk as follows:

\[ Q_{r_{it}}(\tau_j) = \alpha_i(\tau_j) + \beta_i(\tau_j)r_{mt} \quad \text{for all} \quad \tau_i \]

Motivated by the departure of distribution from normality, we propose a weighted average of estimator over multiple quantiles

\[ \hat{\beta}_i = \frac{\sum_{j=1}^{n} \omega_j \beta_i(\tau_j)}{\sum_{j=1}^{n} \omega_j} \quad \omega_j > 0 \quad \sum_{j=1}^{n} \omega_j = 1 \]

The decrease of measurement error for beta risk come from a robust strategy of quantile combination estimation which take advantage of distribution information across multiple quantiles. Therefore we use quantile combination estimator \( \hat{\beta}_i \) as proxy for the true beta risk in the second stage of cross section regression.\(^3\)

\[ r_{it} = \gamma_{0t} + \gamma_{mt}\hat{\beta}_{it} + \epsilon_t \]

Meanwhile, the robust estimation of Quantile Combination can also be conducted in the second pass cross section regression.

\[ \hat{\gamma}_{mt} = \sum_{j=1}^{n} \omega_j \gamma_{mt}(\tau_j) \]

Knez and Ready (1997) found that the risk premium for market beta, size and book to market factor is sensitive to the extreme value of factor loadings in the tail of distribution. Especially, the risk premium on size disappear when the 1 percent most extreme observations are trimmed each month. Whereas, our quantile combination approach may obtain

\(^3\)The proposed robust estimator can be extended to multi-factor model where more than one factor loadings such as size and book to market ratio suggested by Fama French (1993) could be estimated by the robust quantile combination approach as well.
more robust estimator of risk premium without trimming the extreme outliers which may convey useful information

Michelle Barnes and Anthony Hughes (2002) found that, in the small sample estimation, the error in variable (EIV) problem can be relieved by the quantile regression at median\(^4\). Their simulation experiment further demonstrated that, when the EIV problem is rendered more extreme, the bias of the OLS estimator exceeds that of quantile regression for all \(\tau\) under consideration. With this perspective, the methodology of quantile combination gain additional acceptance in the presence of EIV problem since it is nothing more than the weighted average of quantile estimates which may further reduce the bias of estimator.

Furthermore, vast amounts of Financial literatures demonstrate that the risk of beta varies across the entire return distribution. Investors show different behavior in "normal" and extreme market conditions. This differential behavior affects the beta of individual stock, which exhibits time varying property or regime dependence characteristics. Modeling the behavior of CAPM models using quantile regression gains the added advantage of capturing the tail value as well as efficiently analysing the median values. Figure 1(2) presents the average of beta risk over different quantiles for individual firms using monthly(daily) data in our sample, which illustrate the larger impact of market on extreme quantiles of a firms compared to immediate quantiles. Baur and Schulze(2010) exploited investor learning or herding theory to explain this roughly symmetrical U-shape pattern for beta coefficients in the tail quantiles of distribution. Following Patton and Verardo(2009) model, beta is larger if large return surprise(extreme quantiles) occur. Whereas, if news announcements are small or within a "normal range", investors don’t use information from other firms and beta does not increase therefore.

\(^4\)In their paper, they didn’t give further explanation for the reason why quantile regression at median(\(\tau = 0.5\)) outperform least square estimation in the presence of measure error for \(\beta\). We attribute it to the robustness property of quantile regression to extreme outlier deviation(measurement error)
Most of Financial literatures investigate beta risk separately at different quantile. However, quantifying beta risk at a specific quantile of distribution would potentially overestimate or underestimate the overall risk of a stock. In this spirit, combining multiple quantiles over distribution may not merely be motivated by its statistical robustness to outliers in heavy tail, but also due to economic reason that combining quantile would potentially exploit useful information about different causal effect between response and
factors at different part of distribution. In this section, we briefly introduce the motivation to combine multiple quantiles to estimate beta risk. In the next section, we attempt to study how to combine quantiles.

3. Quantile Combination Method

Quantile regression as introduced by Koenker and Bassett (1978) is an extension of the classical least square estimation (OLS). In most of case, the effects of conditional variable on response are not constant, but rather vary across different part of distribution. The least square regression only studies the movement of central tendency of distribution, which may disguise the useful information in the tail or other part of distribution. With the quantile regression, the entire part of distribution can be estimated by choosing a very fine grid of quantile. From this perspective, quantile regression is more powerful than least square regression since it can identify dependence of various part of distribution on conditional variables.

Moreover, the method of quantile regression is characterized by its ”non-parametric” peculiarity. Unlike least square estimation or some other methods, no parametric structure in distribution function needs to be imposed in quantile regression. This advantage wins its wide acceptance in application since the shape of distribution is more often than not unknown to the econometrician.

Furthermore, quantile regression provides a powerful robustness to extreme shocks. It is well established that classic ordinary least square estimation (OLS) will lose its efficiency in the presence of outlier observation or extreme values in the tails of distribution. Quantile regression, is naturally robust to outliers since its cost function is the sum of the absolute values of the residuals, deviant observations are given less importance than that under criteria of least square errors. More generally, an asymmetrically heavy (low) weight is assigned to large positive (negative) deviations.

For the linear regression model
Following Koenker and Basset (1978), the \( \tau \)th conditional quantile regression model can be specified as:

\[
Q_y(\tau|X) = X^T \beta(\tau)
\]

where

\[
\hat{\beta}_\tau = \arg \max_{\beta} \left( \sum_{\epsilon_t > 0} \tau |y_t - X^T \beta| + \sum_{\epsilon_t < 0} (1 - \tau) |y_t - X^T \beta| \right)
\]
combination of sample quantiles. In this section, several combination processes were proposed to investigate how to estimate the weight average to combine multiple-quantiles.

For illustrative purpose, we start from a general linear location-scale shifted model

\[ y_t = \alpha + \beta x_t + \sigma_t \epsilon_t \quad \sigma_t = \gamma_0 + \gamma_1 x_t \]

where \( \epsilon_t \) is IID process and orthogonal to \( x_t \). The conditional quantile regression estimators of parameters can be specified as

\[ \alpha(\tau) = \alpha + \gamma_0 F^{-1}_{\epsilon}(\tau) \quad \beta(\tau) = \beta + \gamma_1 F^{-1}_{\epsilon}(\tau) \]

The presence of heteroskedasticity resulting from the dependence of volatility on co-variate \( x_t \) introduce no asymptotic bias in \( \beta(\tau) \), but it does introduce a source of asymptotic inefficiency. Following Konker and Zhao(1994), we can transform quantile regression model as

\[ \tilde{y}_t = \tilde{\alpha} + \beta \tilde{x}_t + \epsilon_t \]

where \( \tilde{y}_t = y_t / \hat{\sigma}_t \), \( \tilde{x}_t = x_t / \hat{\sigma}_t \). By conducting such model transformation, full efficiency can be achieved since the error term is converted to be IID process in the new model.

In the statistical literatures, one general form of local averaging for Trimmed Regression Quantile(TRQ) combination estimators(Konker and Portnoy(1987), Chan and Lakonishok (1992)) is specified as

\[ \hat{\beta} = \frac{1}{1 - 2\alpha} \int_{\alpha}^{1-\alpha} \hat{\beta}(\tau) d\tau \]

Where \( 0 < \alpha < 0.5 \). All observations lying on or below the \( \alpha \) th quantile (corresponding to large negative outliers), as well as all observations lying on or above the \( (1 - \alpha) \) th quantile (corresponding to large positive outliers) are excluded due to poor estimation in the tail. Each quantile is weighted by its “relative frequency” of occurrence, given by its corresponding interval of percentile value.

If we estimate the parameters of linear factor model at a finite discrete quantiles \( \tau_i \) (i =

\(^5\)To avoid confusion, in the following sections, we still use \( y_t \) to represent \( \tilde{y}_t \), and so on for \( \tilde{x}_t \) unless pointing out explicitly
1, 2, ⋅⋅⋅, K), then the weighted average quantile based estimators\(^6\) can be constructed as

\[
\hat{\beta} = \sum_{i=1}^{k} \omega_i \hat{\beta}(\tau_i) \quad \omega_i > 0, \quad \sum_{i} \omega_i = 1
\]

The simplest and most "naive" quantile combination estimation: Mean Quantile Combination (MQC) is equally weighted average estimator

\[
\hat{\beta} = \frac{1}{T} \sum_{i=1}^{T} \hat{\beta}(\tau_i)
\]

It turns out that such simple average of multiple quantiles, sometimes perform best since it avoid to estimate combination weight which may induce measurement errors in the model.

In addition, one widely used simple linear combination of regression quantile is the Tukey trimean which assign predetermined weights to three quantiles from left tail to right tail

\[
\hat{\beta}_{TRM} = 0.25 \hat{\beta}_{0.25} + 0.5 \hat{\beta}_{0.5} + 0.25 \hat{\beta}_{0.75}
\]

Whereas, all of the above proposed combinations of quantiles are not real data driven and very easy to implement because the weight of each quantile is predetermined. Whereas, we proposed a real data dependent quantile combination where the weight is estimated by the conditional density of the target variable

Described by the Konker(2005), consistency and asymptotic normality of quantile coefficients can be established under certain mild conditions

\[
\sqrt{n} \left( \hat{\beta}(\tau) - \beta(\tau) \right) \sim N(0, \tau(1-\tau)D_1^{-1}D_0D_1^{-1})
\]

where

\[
D_0 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T \quad D_1(\tau) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f_{\epsilon}(F_{\epsilon}^{-1}(\tau)) x_i x_i^T
\]

In the IID error model

\[
\sqrt{n} \left( \hat{\beta}(\tau) - \beta(\tau) \right) \sim N \left( 0, \frac{\tau(1-\tau)}{f_{\epsilon}^2(F_{\epsilon}^{-1}(\tau))} D_0 \right)
\]

\(^6\)Weighted average quantile regression belongs to the class of L-estimators, which are obtained as linear combination of order statistics(Konker(2005))
The quantile regression estimator of $\beta$ at $\tau$ has an asymptotic variance matrix in the form as

$$
\frac{\tau(1-\tau)}{f_\varepsilon^2(F^{-1}(\tau))} D_0
$$

The precision of $\beta$ estimator at each quantile $\tau_i$, therefore, can be measured by its inverse asymptotic variance. A higher value of asymptotic variance represents a lower precision of $\beta_{\tau_i}$. The major source of quantile heteroskedasticity is only reflected by the scalar term

$$
\frac{\tau(1-\tau)}{f_\varepsilon^2(F^{-1}(\tau))}
$$

Following Xiao(2011), therefore, we can weighted each quantile estimation by the standarized value of this scalar term :

$$
\omega_i = \frac{f_\varepsilon(F^{-1}(\tau_i))}{\sqrt{\tau_i(1-\tau_i)}} / S_w \quad \text{where} \quad S_w = \sum_i \frac{f_\varepsilon(F^{-1}(\tau_i))}{\sqrt{\tau_i(1-\tau_i)}} \quad (4)
$$

Note that the weight average $\omega_i$ now is not a predetermined constant but instead depends on the shape of distribution, which could be adjusted automatically based on the estimated density function over various quantiles. This is important when the distribution becomes skewed where the density is no longer symmetrical.

Given a family of estimated conditional quantile function, it is straightforward to estimate the conditional density of residuals $f_\varepsilon(F^{-1}(\tau_i))$ as (see Koenker 2005)$^7$:

$$
f_\varepsilon(F^{-1}(\tau_i)) = \frac{2h}{F^{-1}(\tau_i + h) - F^{-1}(\tau_i - h)} \quad (5)
$$

where $h$ is the bandwidth$^8$ which connect two neighboring points on the cumulative probability function.

Since the measurement of probability density function at extreme tail distribution could be imprecise due to lack of observation, the information in the extreme tail would be discarded by using the truncated version of weight average. More generally, given a small number of $\xi$, i.e $\xi = f_\varepsilon(F^{-1}(0.05))$, we can discard information in tail below 5% quantile or

---

$^7$The conditional densities can also be estimated by the Epanechnikov kernel, which would generate smooth densities when the time series sample size is short.

$^8$Bandwidth for each quantile can be estimated by Bofinger(1975) or Hall & Sheather(1988) method where $h_n = \frac{z_\alpha^{2/3}[1.5\phi^2(\Phi^{-1}(\tau))]/(2(\Phi^{-1}(\tau))^2 + 1)^{1/3}]n^{1/3}}$ and $z_\alpha$ satisfies $\Phi(z_\alpha) = 1 - \alpha/2$.
upper 95\% quantile by multiplying the indicator function $I(|f_\epsilon(F^{-1}_\epsilon(\tau_i)) > \xi)$

$$\omega_i = \frac{f_\epsilon(F^{-1}_\epsilon(\tau_i))I(|f_\epsilon(F^{-1}_\epsilon(\tau_i)) > \xi)}{\sqrt{\tau_i(1-\tau_i)}} / S_w$$

where $S_w = \sum \frac{f_\epsilon(F^{-1}_\epsilon(\tau_i))I(|f_\epsilon(F^{-1}_\epsilon(\tau_i)) > \xi)}{\sqrt{\tau_i(1-\tau_i)}}$

(6)
4. Monte Carlo Simulation and Estimation Performance

We use Monte Carlo simulation to access the impact of tail thickness or skewness of error distribution on the relative efficiency of the various regression estimation on beta risk. Following the simulation exercise of McDonald, Michelfelder & Theodossiou (2009) and among others, a simulated series for the stock rate of return is generated by a specific model following CAPM

\[ R_t = \alpha + \beta R_{M,t} + \epsilon_t \]

where \( R_{M,t} \) is the excess market return for entire period of sample data. The true value of parameter \( \alpha \) and \( \beta \) are set to be zero and one, respectively since we assume that the CAPM model hold in the simulation.

Considering the empirical evidence that the distribution of finance time series departs from normality and is characterized by thick tails and skewness, the regression errors \( \epsilon_t \) are generated by student t, skewed student t and benchmark distribution of normality. We simulate student t distribution with various degree freedom to examine if the degree of tail thickness affects estimation performance. A family of Skewed student t distribution (Hansen 1994; Fernandez and Steel 1998; Lambert and Laurent 2009) are generated as well to account for the effects of both skewness and tail thickness. We consider two skewed \( t(4) \) distribution with skewness set to be 0.5 (lower skewed) and 0.8 (higher skewed) \(^9\)

To ensure our simulation exercise more closely approximate the actual distribution of stock market return, in each simulation, we use 534 observations on the excess market return over sample period from July 1964 to December 2008 to proxy for market excess return \( R_{M,t} \). Residuals \( \epsilon_t \) are drawn from the above hypothesized distributions. The gen-

\(^9\) As matter of fact, we simulate various skewed and fat tailed distribution like mixed-normal(thick tailed variance contaminated), skewed log normal and standard inverse Gaussian distribution, the estimation results, however, turn out not to be quite different. To save space, we only report the simulation result for \( t \) family distribution since it is more convenient to generate different shape of distribution with various skewness and thick tail.
erated stock return \( r_t \) is then regressed on the market risk premium \( r_{m,t} \) using various estimation strategies: Ordinary Least Square(OLS), Least Absolute Deviation(LAD) and the various Quantile combination strategies proposed in the section 2.

As proposal estimation methodology is capable of utilizing distribution information across multiple quantiles, we construct our quantile combination estimation by exploiting the distribution information evenly from left tail through central to right tail of distribution. Ten equally divided percentiles \( \tau = (0.1, 0.2, 0.3, \cdots, 0.9) \) are selected.\(^{10}\)

For each simulation process, a total of 5000 replications are carried out. We use the ratio of root mean square error(RMSE) to evaluate the relative accurateness of parameter estimation from various methods. Three cases of simulation are considered in this section. In table 1, the distribution of error terms are IID process without conditional heteroscedasticity. Quantile-based estimation outperform OLS estimation without exception even when the distribution is asymmetrical (positive or negative skewed). In Table 2, we introduce heteroskedasticity component in the distribution of error term. Thus the model is the linear location and scale shifted model. And quantile combination estimation are carried out without making any correction. It turns out that quantile combination would lose its strength in the presence of both conditional heteroskedasticity and skewness. In table 3, we show that the robustness of quantile combination method can be revived by correcting the heteroskedasticity components and transforming it to be homoskedasticity linear model by dividing the estimated volatility on the both sides of linear model.

\(^{10}\)Theoretically, more values of quantiles combined represents more distribution information incorporated. However, there is trade off whether combine more quantile or not. More quantiles included may induce the problem of quantiles crossing, which need imposing non-crossing restriction on quantiles. If the model is misspecified, more quantile combination may introduce more error in the model. Moreover, more quantiles to be combined would call for more computation.
Table 1

Performance Evaluation for the Estimated Slope $\hat{\beta}$ from CAPM Model

\[ R_{i,t} = \alpha + \beta R_{m,t} + \epsilon_t \]

Ratio of RMSE : \[ \frac{V_{ols}}{V_i} \]

Where \[ V_i = \left( \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta)^2 \right) \] and \[ \hat{\beta} = \sum_{j=1}^{k} \omega_j \hat{\beta}(\tau_j) \]

<table>
<thead>
<tr>
<th></th>
<th>$\epsilon_t \sim \text{Normal}$</th>
<th>$\epsilon_t \sim t_3$</th>
<th>$\epsilon_t \sim t_5$</th>
<th>$\epsilon_t \sim t_8$</th>
<th>$\epsilon_t \sim \text{SkewT}(4,0.5)$</th>
<th>$\epsilon_t \sim \text{SkewT}(4,0.8)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Median</td>
<td>0.78</td>
<td>1.23</td>
<td>0.96</td>
<td>0.88</td>
<td>1.04</td>
<td>1.03</td>
</tr>
<tr>
<td>Trimean</td>
<td>0.91</td>
<td>1.33</td>
<td>1.07</td>
<td>1</td>
<td>1.19</td>
<td>1.15</td>
</tr>
<tr>
<td>MQC</td>
<td>0.99</td>
<td>1.21</td>
<td>1.06</td>
<td>1.03</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td>QCI</td>
<td>0.98</td>
<td>1.32</td>
<td>1.10</td>
<td>1.04</td>
<td>1.35</td>
<td>1.59</td>
</tr>
<tr>
<td>QCII</td>
<td>0.92</td>
<td>1.36</td>
<td>1.09</td>
<td>1.02</td>
<td>1.39</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Monte carlo simulation based on 5000 replication of simulated data with 534 observation per replication for the sample data between 196407 and 200812. Simulated data based on : $\alpha = 0$; $\beta = 1$; $\tau = [0.1, 0.2, 0.3, \cdots, 0.9]$. In column (2)-(7), $\epsilon_t$ follow standard normal, $t$ distribution with $df = 3, 5, 8$ and Skewed Student $t$ distribution with degree freedom $df = 4$ and skewness parameter $\lambda = 0.8$ or $\lambda = 0.5$

1 MQC refer to simple average of equal weighted quantile combination $\hat{\beta} = \frac{1}{T} \sum_{i=1}^{T} \hat{\beta}(\tau_i)$

2 Trimean indicate Tukey’s trimean estimation $\hat{\beta} = 0.25\hat{\beta}(0.25) + 0.50\hat{\beta}(0.50) + 0.25\hat{\beta}(0.75)$

3 QCI is the density weighted quantile combination where the weight is estimated as : $\omega_i = \frac{f_{\hat{\epsilon}}(F_{\epsilon}^{-1}(\tau_i)) I(|f_{\hat{\epsilon}}(F_{\epsilon}^{-1}(\tau_i))|>\xi)}{\sqrt{\tau_i(1-\tau_i)}} / S_w$

4 QCII is the truncated version density weighted quantile combination where the weight is estimated as : $\omega_i = \frac{f_{\hat{\epsilon}}(F_{\epsilon}^{-1}(\tau_i)) I(|f_{\hat{\epsilon}}(F_{\epsilon}^{-1}(\tau_i))|>\xi)}{\sqrt{\tau_i(1-\tau_i)}} / S_w$

Table I show the performance evaluation for the estimated slope $\beta$ from CAPM model without conditional heteroskedasticity. The estimation results reveal Some facts deserved to be highlighted. Firstly, the efficient gain for robust estimator of $\beta$ is closely related with tail thickness. In the extreme case where error term $\epsilon_t$ are drawn from the symmetric heavy tail $t$ distribution with $df = 3$, quantile combination estimation reduce measurement error significantly (The largest $\frac{RMSE_{ols}}{RMSE_{\epsilon_t}} = 1.36$). This is a substantial improvement for
the measurement of beta risk or cost of capital. Although Median regression (quantile at 50%) outperform OLS estimation in heavy tail distribution, combing multiple quantiles can further decrease measure error of estimators. As the degree of tail thickness become less (the degree of freedom increase from 3 to 8), the efficiency improvement become smaller as well, but it still reveals out-performance of quantile combination estimator relative to OLS under mild fat tail distribution. Column (2) to (4) show that, when the data follow a symmetric heavy tail distribution, combining multiple quantiles can substantially improve efficiency of estimation for slope coefficient $\beta$ in linear regression model.

Secondly, even in the presence of asymmetry of distribution (positive or negative skewness), Quantile combination approach can still yield more efficient estimator. Without conditional heteroskedasticity, the efficiency gain for $\beta$ estimation is still pronounced from quantile combination where $\beta(\tau) = \beta$ is not contaminated by residual quantile $Q_{\epsilon(\tau)}$ in the location shift model. Following Hansen (1994), Fernandez and Steel 1998, Lambert and Laurent 2009 and among others, We simulate skewed heavy tail t distribution with different degree of skewness. Column (6) and (7) reveal that beta $\beta$ estimation can be substantially improved from quantile combination under the heavy tailed distribution, even though the distribution is highly positive skewed. Moreover, the density weighted quantile combination perform much better than the predetermined weighted combination since the former weight is data-driven which can be adjusted depending on the shape of distribution.

The first column reported in the table 1 assumes the "ideal" situation where the CAPM model residuals $\epsilon_t$ are drawn from normal distribution with mean zero and standard deviation 11.66% \(^{11}\) such that a typical average $R^2 = 19.66$ percent is yielded for OLS estimation for the data between 196407 to 200812. In this setting, OLS estimation perform best as expected. However, the efficiency loss for the quantile combination is trivial for $\beta$ estimation. Combining multiple quantiles yields estimators very close to OLS estimators. This is a good news since the distribution of financial data is usually unknown to econometrician, the cost of applying quantile combination is trivial even in the "worst

\[^{11}\] A typical linear regression model shows that $R^2 = 1 - \frac{\sigma^2 Var(r_{M,t})}{\sigma^2}$, which leads to $\sigma^2 = \frac{1-R^2}{R^2} \beta^2 Var(r_{M,t})$. Therefore, we simulate $\epsilon \sim N(0, 0.1166)$ such that a typical average $R^2 = 19.66$ percent is yielded for OLS estimation of CAPM model in the sample data between 196407 and 200812.
scenario” when the true distribution exactly follow normal distribution

Table 2
Performance Evaluation for the Estimated Slope \( \hat{\beta} \) from CAPM Model with Conditional Heteroskedasticity

\[
R_{i,t} = \alpha + \beta R_{m,t} + \sigma_t \epsilon_t \quad \text{Where} \quad \sigma_t = \gamma_0 + \gamma_1 R_{m,t}
\]

Ratio of RMSE: \( \frac{V_{ols}}{V_i} \) Where \( V_i = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (\hat{\beta}_t - \beta)^2} \) and \( \hat{\beta} = \sum_{j=1}^{k} \omega_j \hat{\beta}(\tau_j) \)

<table>
<thead>
<tr>
<th></th>
<th>( \epsilon_t \sim \text{Normal} )</th>
<th>( \epsilon_t \sim t_3 )</th>
<th>( \epsilon_t \sim t_5 )</th>
<th>( \epsilon_t \sim t_8 )</th>
<th>( \epsilon_t \sim \text{SkewT}(4,0.5) )</th>
<th>( \epsilon_t \sim \text{SkewT}(4,0.8) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Median</td>
<td>1</td>
<td>1.08</td>
<td>1.72</td>
<td>1.33</td>
<td>1.21</td>
<td>0.38</td>
</tr>
<tr>
<td>Trimean</td>
<td>1.12</td>
<td>1.76</td>
<td>1.37</td>
<td>1.25</td>
<td>0.41</td>
<td>0.32</td>
</tr>
<tr>
<td>MQC</td>
<td>1.14</td>
<td>1.70</td>
<td>1.34</td>
<td>1.24</td>
<td>0.49</td>
<td>0.38</td>
</tr>
<tr>
<td>QCI</td>
<td>1.14</td>
<td>1.76</td>
<td>1.37</td>
<td>1.26</td>
<td>0.43</td>
<td>0.33</td>
</tr>
<tr>
<td>QCII</td>
<td>1.11</td>
<td>1.76</td>
<td>1.36</td>
<td>1.24</td>
<td>0.39</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Monte Carlo simulation based on 5000 replication of simulated data with 534 observation per replication for the sample data between 196407 and 200812. Simulated data based on: \( \alpha = 0; \beta = 1; \tau = [0.1, 0.2, 0.3, \ldots, 0.9] \). In column (2)-(7), \( \epsilon_t \) follow standard normal, t distribution with \( df = 3, 5, 8 \) and Skewed Student t distribution with degree freedom \( df = 4 \) and skewness parameter \( \lambda = 0.8 \) or \( \lambda = 0.5 \)

1 MQC refer to simple average of equal weighted quantile combination \( \hat{\beta} = \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}(\tau_t) \)
2 Trimean indicate Tukey’s trimean estimation \( \hat{\beta} = 0.25\hat{\beta}(0.25) + 0.50\hat{\beta}(0.50) + 0.25\hat{\beta}(0.75) \)
3 QCI is the density weighted quantile combination where the weight is estimated as: \( \omega_i = \frac{f_i(F^{-1}(\tau_i))}{\sqrt{\tau_i(1-\tau_i)}} / S_w \)
4 QCII is the truncated version density weighted quantile combination where the weight is estimated as: \( \omega_i = \frac{f_i(F^{-1}(\tau_i))I(\|f_i(F^{-1}(\tau_i))\|<\xi)}{\sqrt{\tau_i(1-\tau_i)}} / S_w \)

So far in the simulation exercise, we assume that the beta risk is constant over the whole distribution. But most of Financial literatures have shown that the evaluation of downside risk is quite different from upside risk. The more commonly
used model is location and scale shift linear model where both the intercept and slope coefficients change with quantiles. We constructed a conditional heteroskedasticity linear factor model where the volatility of residuals is simple AR(1) model. Therefore the quantile estimators of both intercept and slope coefficients contain a component of the quantile of residuals

\[
\alpha(\tau) = \alpha + \gamma_0 F^{-1}_\epsilon(\tau) \quad \beta(\tau) = \beta + \gamma_1 F^{-1}_\epsilon(\tau)
\]

Table 2 evaluates the performance of different estimations in the presence of conditional heteroskedasticity. Column (2) to (5) present the lower value of mean square error for quantile combination estimators of \(\beta\) when distribution is symmetrical and heavily tailed. It is surprising that quantile-based estimation also outperform OLS estimation even in the 'ideal' situation when residuals \(\epsilon_t\) follows normal distribution. However, when the distribution is characterized by both asymmetry and conditional heteroscedasticity, the benefit of quantile combination would disappear, and even deteriorates the performance of quantile combination estimations. Column (6) and (7) reveal greater measurement error for the slope coefficient \(\beta\) for all quantile combination estimators without exception. In other word, in the presence of both conditional heteroskedasticity and skewness, Quantile combination approach seems to lose its strength to yield more efficient estimator.
Table 3
Performance Evaluation for the Estimated Slope $\hat{\beta}$ from transformed CAPM Model with Conditional Heteroskedasticity and Skewness

\[ R_{i,t} = \alpha + \beta R_{m,t} + \sigma_t \epsilon_t \quad \hat{\sigma}_t = \hat{\gamma}_0 + \hat{\gamma}_1 R_{m,t} \]

\[ R^*_i,t = \alpha^* + \beta^* R^*_m,t + \sigma_t \epsilon_t \quad \text{where} \quad R^*_i,t = \frac{R_{i,t}}{\hat{\sigma}_t}, \quad R^*_m,t = \frac{R_{m,t}}{\hat{\sigma}_t}, \quad \alpha^* = \frac{\alpha}{\hat{\sigma}_t} \]

<table>
<thead>
<tr>
<th></th>
<th>$\epsilon_t \sim \text{SkewT}(4, 0.2)$</th>
<th>$\epsilon_t \sim \text{SkewT}(4, 0.5)$</th>
<th>$\epsilon_t \sim \text{SkewT}(4, 0.8)$</th>
<th>$\epsilon_t \sim \text{SkewT}(4, 1.2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Median</td>
<td>1.89</td>
<td>2.47</td>
<td>3.09</td>
<td>2.13</td>
</tr>
<tr>
<td>Trimean</td>
<td>1.97</td>
<td>2.58</td>
<td>3.17</td>
<td>2.24</td>
</tr>
<tr>
<td>MQC</td>
<td>2.04</td>
<td>2.70</td>
<td>3.28</td>
<td>2.36</td>
</tr>
<tr>
<td>QCI</td>
<td>1.50</td>
<td>1.75</td>
<td>1.75</td>
<td>1.71</td>
</tr>
<tr>
<td>QCII</td>
<td>1.47</td>
<td>1.72</td>
<td>1.72</td>
<td>1.64</td>
</tr>
</tbody>
</table>

Finally, we construct a simulation experiment to investigate how to conduct quantile combination when the distribution is featured by both conditional heteroskedasticity and asymmetry. Table 3 documents the simulation result of quantile combination for the transformed CAPM model. We simulated four skewed t distribution with different skewness parameters to study the impact of skewness on the robustness of our quantile-based model. Column (2) to (5) reveal that the advantages afforded by quantile combination revives after transforming the linear CAPM model by dividing its estimated volatility\(^{12}\), among which, the simple mean average of quantile combination estimation perform best since it avoid to estimate weights which may introduce measurement error in the model.

In summary, our simulation experiment reveal that Quantile combination estimation can produce robust and more efficient estimators for beta risk under the heavy tail distribution. In the presence of skewness but not conditional heteroskedasticity, Quantile Combination can still yields robust estimators for slope coefficient

\(^{12}\text{The volatility can be estimated using a stochastic volatility or GARCH model if distribution assumption of } \epsilon \text{ are known. Here we estimate volatility by a simple AR(1) model}\)
Lastly, with the presence of both skewness and conditional heteroskedasticity, robustness of quantile combination method seem to lose its strength. However, transforming the conditional heteroskedasticity model to be homogenous model by dividing the estimated volatility, quantile combination approach revive its strength to produce more efficient estimator for beta risk.
5. Empirical Studies

In this section, we start to investigate if our robust estimation strategy can really produce more efficient estimator for beta risk in the real data instead of only in simulation experiment. We estimate the individual firms’ beta by various quantile combination approach and examine if asset pricing test can be improved by our advocated more efficiently estimated beta.

5.1. Data

We consider a population of stocks traded on the NYSE, AMEX and NASDAQ from July of 1964 to the December of 2008. Any stock was removed that did not have at least 120 months (or 2520 trading days) of data during the sample period. We don’t remove any stocks that de-listed or stop trading due to liquidity to avoid survival bias. In line with Avramov and Chordia(2006), we only includes the stocks that have data available in month $t - 1$ to compute the firm characteristics size, book to market ratio. Following Fama and French (1993), we exclude firms with negative book to market equity. Imposing this restriction results in a universal of 5381 firms over the sample period.

Many empirical studies had suggested the use of high frequency sample interval, say daily or intra-daily data in order to estimate beta risk more accurately. On the one hand, the distribution of daily stock returns exhibit much fatter tails than monthly stock returns. Our robust quantile-based estimation is expected to perform better with thicker tail distribution. On the other hand, the distribution of daily stock return is characterized as well by the higher skewness. Monte Carlo simulation above reveal that the efficiency gain of quantile combination estimation would be hurt or diluted by the highly skewed distribution of residuals. Table 4 below present the test of normality assumption for Monthly and Daily stock return in our sample. Obviously, the departure of normality is much more serious for daily data than monthly data. Daily stock return exhibit more significant and higher value of skewness and excess kurtosis overall.\footnote{The departure of distribution from normality also depends on the length of window of time series that was investigated. In the short window span of data, the distribution is close to normality even for high}
In order to capture the impact of tail thickness but avoid the disturbance of skewness effect, we intentionally conduct a data manipulation by randomly selecting about 670 firms in our sample that have a high and significant value of excess kurtosis, but low value of skewness for return distribution\textsuperscript{14}. Then we analyze how systematic $\beta$ risk could be assessed more accurately by our quantile-based combination approach in the presence of fat tail but nor very skewed distribution.

### Table 4
Test of normality assumption for Monthly and Daily data (196407 - 200812)

The normality assumption is tested for monthly and daily stock return in our sample between July 1964 and December 2008. In each panel, the column labeled "Average Value" report the average value of skewness and excess kurtosis for the security return in our daily or monthly population. The column labeled "$\alpha = 1\%$" refers to the percentage of stocks for which the null hypothesis of zero skewness, or zero excess kurtosis, or Jarque Bera Normality assumption is rejected at the 1% significant levels, respectively. And vice versa for the column $\alpha = 5\%$

<table>
<thead>
<tr>
<th></th>
<th>Monthly Data</th>
<th></th>
<th></th>
<th>Daily Data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Value</td>
<td>$\alpha = 1%$</td>
<td>$\alpha = 5%$</td>
<td>Average Value</td>
<td>$\alpha = 1%$</td>
<td>$\alpha = 5%$</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.8951</td>
<td>0.6720</td>
<td>0.7475</td>
<td>0.9843</td>
<td>0.9277</td>
<td>0.9467</td>
</tr>
<tr>
<td>exKurt</td>
<td>5.8504</td>
<td>0.9214</td>
<td>0.9397</td>
<td>23.6048</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Normality</td>
<td>0.8659</td>
<td>0.9294</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

We present the estimation of beta risk from different methods. Figure 3(4) show the distribution of the difference between OLS estimator and various quantile combination estimator of $\beta$ risk using monthly(daily) data. To save the space, only 4 different quantile combination estimators are compared with OLS estimators, among which, we define $\beta_{MQC}$ as simple average of quantile combination, $\beta_{QC1}$

\textsuperscript{14}We select firms with return that have significantly positive excess kurtosis, but low value of skewness between -1 and 1, which results in a universal of 671 firms in our sample
refers to density weighted quantile combination, $\beta_{\text{Trimean}}$ is Tukey trimean quantile combination and $\beta_{\text{TRQ}}$ is defined as trimmed regression quantile (TRQ), all of which have been discussed in section 2.

In Figure 3 we observe some difference between OLS estimators and various quantile combination estimators. The difference of two estimators even exceed 0.5 for some of firms, which might imply a large efficient loss for OLS estimation of beta risk, given that quantile combination approach yield more robust estimator under the fat tail distribution. Moreover, the distribution of two estimators difference seems to skewed slightly toward left tail, which suggest that the OLS estimator is more likely to overestimate the true value of beta risk. As we have presented in the table 4, the distribution of daily data exhibit much thicker tail than monthly data, which reflect larger deviation between OLS estimator and our robust quantile-based estimator in Figure 4. It is surprising that these differences are all skewed toward right tails without exception, which implies that OLS estimator in daily data is more likely to underestimate the true value of beta coefficient. At the current stage, we haven’t accounted for the time variability of beta coefficient. This may conceal the true difference between OLS and our robust estimators.
Since the true value of beta is unobservable, it is impossible to conduct a real out of sample predictability analysis\textsuperscript{15}. However, we can test the precision of beta estimation indirectly by the widely used Fama-MacBeth two pass asset pricing test of CAPM. Be aware that the goal of this paper is not to verify the validity of CAPM model. Instead, we want to test if beta risk estimated by quantile combination approach outperform OLS estimator in the asset pricing test. The test is motivated by the intuition that the higher precision in estimating latent variable betas may lead to more powerful cross sectional stock regression, even though the reliability of this test depends on the validity of CAPM model.

Following the Fama and MacBeth two-pass Methodology\textsuperscript{(1973)}, a stock’s beta risk can be estimated by an ordinary lease square estimation. Instead, we estimate

\textsuperscript{15}Mathijs Cosemans et al (2010) construct a Out-of-Sample Beta Forecast by using the realized beta as proxy for true beta , which is estimated from high frequency data. Because of microstructure friction, the realized beta estimation is only valid for liquid stock.
quantiles of the entire conditional distribution. The model can be written as follows:

\[ Q_{r_{it}}(\tau_j) = \alpha_i(\tau_j) + \beta_i(\tau_j) r_{mt} \text{ for all } \tau_i \]

### Table 5

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>MQC</th>
<th>QCI</th>
<th>QCII</th>
<th>Trimean</th>
<th>TRQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_0 )</td>
<td>0.11</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>( t_{stat} )</td>
<td>0.5013</td>
<td>0.6048</td>
<td>0.5984</td>
<td>0.6023</td>
<td>0.6312</td>
<td>0.6443</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>0.61</td>
<td>0.64</td>
<td>0.65</td>
<td>0.66</td>
<td>0.65</td>
<td>0.59</td>
</tr>
<tr>
<td>( t_{stat} )</td>
<td>1.9964</td>
<td>1.9959</td>
<td>2.0131</td>
<td>2.0297</td>
<td>1.9772</td>
<td>2.0100</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>9.43</td>
<td>9.72</td>
<td>9.69</td>
<td>9.65</td>
<td>9.57</td>
<td>9.53</td>
</tr>
</tbody>
</table>

MQC stands for using simple mean average quantile combination(\( \beta = \frac{1}{n} \sum_{i=1}^{n} \beta(\tau_i) \)) estimator of \( \beta \) in second pass of cross sectional regression. QCI refer to density weighted quantile combination. QCII stands for truncated version of QCI with truncated quantile \( \tau = [0.2, 0.3, \ldots, 0.80] \). Trimean is Tukey Trimean quantile combination. TRQ is trimmed regression quantile with \( \alpha = 0.20 \)

Neweywest t statistics are reported for each parameter.

Therefore, we propose a weighted average of estimator over multiple quantiles as proxy for the true beta risk in the second stage of cross section regression at each time \( t \)

\[ \hat{\beta}_i = \sum_{j=1}^{n} \omega_j \beta_i(\tau_j) \]

\[ r_{it} = \gamma_0 t + \gamma_{mt} \hat{\beta}_it + \epsilon_t \]

Then we compute time series average of the cross sectional regression estimates

\[ \gamma_0 = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{0,t} \quad \gamma_1 = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{m,t} \]

Where \( \gamma_0 \) is the return unexplained by systematic risk. \( \gamma_{mt} \) is the market risk premium. A reasonable estimator of \( \beta \) might produce a positive and significant
estimator of market risk premium. CAPM model predict that the variation of cross sectional stock return can be explained completely by market risk exposure

\[ Er_i = \gamma_m \beta_i \]

Insignificance of the intercept \( \gamma_0 \) from zero indicates a good fitness of CAPM model. Under fat tail distribution, our quantile combination approach is expected to yield more efficient estimates of beta risk with lower measurement error, which would improve the goodness fit of the cross sectional regression.

Table 5 report the Fama MacBeth second pass regression coefficients using beta estimators from different estimation methods. We can see that the goodness of fit \( R^2 \) for the second pass cross sectional regression increase slightly as we use quantile combination approach instead of OLS to estimate beta, which implies a lower measurement error of beta risk in the first pass time series regression, leading to the improvement in the second pass cross section regression. The estimator of market risk premium \( \gamma_1 \) are positive and significant at a level of 5%. And the abnormal return \( \alpha \) are all insignificant, indicating the reasonable performance of CAPM model in our sub-sample data, even without other factors that have typically help to resuscitate the CAPM. It is not surprising that the improvement of cross sectional regression is trivial and not impressive, considering that all estimation of beta ignore its time varying dynamics. Even though our robust estimator expect to improve the efficiency of beta risk under heavy tail distribution, this efficiency improvement can be concealed by the misspecification of CAPM model due to its failing to account for instability of beta. Therefore in the following section, we introduce time varying characteristics of beta in our quantile combination estimation.

5.2.1 Rolling Forecast Test.

It is widely accepted that beta risk is not constant, but instead time varying over business cycle. Modeling time varying behavior of beta risk has been focus of interest for the studies of CAPM model over the past few decades. One of most widely
used model estimating time varying conditional beta is based on purely data-driven rolling sample methodology, which can be dated back to Fama and MacBeth(1973) who applied OLS estimation sequently to a local window length of data to estimate parameters($\alpha_{it}, \beta_{it}$) of CAPM models.

\[ R_{it} = \alpha_{it} + \beta_{it} R_{mt} + u_{it} \]

Most empirical studies, following Fama and MacBeth(1973), use a $r = 60$ observation window for monthly data to estimate time changing beta. Given the non-normality distribution of stock return, we apply our quantile-based approach in each window to estimate beta risk. As we pointed out before, the goal of this paper is not to test validity of CAPM model, neither does it intend to judge whether the rolling estimation is adequate to model time varying behavior of beta risk. Instead, we attempt to show if quantile combination approach produce more efficient estimator of beta risk under fat tailed distribution, assuming that the rolling window estimation can capture or at least partially capture dynamics of beta risk.

We introduce a hybrid estimation strategy for our robust quantile combination approach in the rolling window framework. In each short window length of data(60 months), the departure of distribution from normality was test firstly. If the evidence of non-normality and excess kurtosis is statistically significant, quantile combination approach was implemented to obtain a robust estimates of beta risk. Otherwise, OLS estimation remains to estimate beta. Therefore, we may sequently switch our estimation strategies from quantile combination approach to OLS in each local window of data, depending on the deviation of distribution from normality is statistically prominent or not. Such hybrid estimation strategy presumably enable us to obtain efficient estimates of beta risk dynamically.
Table 6
Fama MacBeth second pass regression coefficients for monthly sample data (196407-200812)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>MQC</th>
<th>QCI</th>
<th>QCII</th>
<th>Trimean</th>
<th>TRQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>0.18</td>
<td>0.20</td>
<td>0.22</td>
<td>0.22</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>$t_{stat}$</td>
<td>(0.6834)</td>
<td>(0.7729)</td>
<td>(0.8479)</td>
<td>(0.8718)</td>
<td>(0.9066)</td>
<td>(0.9569)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.48</td>
<td>0.49</td>
<td>0.48</td>
<td>0.0.48</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>$t_{stat}$</td>
<td>(1.5995)</td>
<td>(1.6095)</td>
<td>(1.5701)</td>
<td>(1.5653)</td>
<td>(1.4836)</td>
<td>(1.4827)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>7.89</td>
<td>8.07</td>
<td>8.04</td>
<td>7.95</td>
<td>7.79</td>
<td>8.00</td>
</tr>
</tbody>
</table>

MQC stands for using simple mean average quantile combination ($\beta = \frac{1}{n} \sum^n_i \beta(\tau_i)$) estimator of $\beta$ in second pass of cross sectional regression. QCI refer to density weighted quantile combination. QCII stands for truncated version of QCI with truncated quantile $\tau = [0.2, 0.3, \cdots, 0.80]$. Trimean is Tukey Trimean quantile combination. TRQ is trimmed regression quantile with $\alpha = 0.20$

Neweywest t statistics are reported for each parameter.

Table 6 presents the Fama MacBeth second pass regression coefficients in the rolling window framework. Undoubtedly, the performance of CAPM model is not as good as expected since the market risk premium $\gamma_1$ is insignificant from zero at 5% level. The model goodness of fit $R^2$ declines compared with that in table 5, indicating that rolling window estimation fails to capture the true time varying behavior of beta risk. However, the goodness of fit $R^2$ for second pass cross sectional regression still increase slightly as we use quantile combination approach instead of OLS to estimate beta, which implies a lower measurement error of beta risk in the first pass time series regression, leading to the improvement in the second pass cross section regression.

5.3 Robust measurement of Factor Premium.
Since Fama and French (1992,1993), there has been a vigorous and ongoing debate on whether size and book to market factors are really priced in explaining cross sectional stock return. Knez and Ready (1997) found that the size and book to market risk premium is sensitive to the influential observation (outlier). By trimming 1 percent of most extreme observation each month, they found that the
size effect almost disappear. However, they also conceded that trimming extreme outliers may not be appropriate since it discards useful information contained in the tail of distribution. Whereas, quantile regression provide a natural tool to analyze the impact of extreme outliers without trimming the influential and potentially informative observations. Following Allen and Singh(2010), we applied quantile combination approach in the second pass of cross section regression to yield a more robust measurement of risk premium across quantiles.

Table 7
Result of Robust estimation for Fama MecBeth Cross section Regression

<table>
<thead>
<tr>
<th>OLS</th>
<th>MQC</th>
<th>QCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β log(ME) log(BE/ME)</td>
<td>β log(ME) log(BE/ME)</td>
</tr>
<tr>
<td>β</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.62</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>(2.0313)</td>
<td>(1.8129)</td>
<td>(1.7071)</td>
</tr>
<tr>
<td>-0.09</td>
<td>-0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>(-2.6958)</td>
<td>(-0.5007)</td>
<td>(1.0223)</td>
</tr>
<tr>
<td>0.13</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>(1.9378)</td>
<td>(1.5604)</td>
<td>(1.1412)</td>
</tr>
<tr>
<td>0.77</td>
<td>-0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>(2.9318)</td>
<td>(-2.5901)</td>
<td>(0.19481)</td>
</tr>
<tr>
<td></td>
<td>0.49</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(1.9742)</td>
<td>(0.48335)</td>
</tr>
<tr>
<td></td>
<td>-0.00</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(-0.0279)</td>
<td>(1.5608)</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.19481)</td>
<td>(1.3935)</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.3619)</td>
<td></td>
</tr>
</tbody>
</table>

The testing procedure can be proceeded as following:
1. For each time period t, the quantile regression model for Fama MecBeth cross section regression was conducted as

\[ Q_r(\tau) = \gamma_0(\tau) + \gamma_{mt}(\tau)\beta \quad \text{for all } \tau = (0.1, 0.2, \cdots, 0.9) \]

2. Implement robust estimation for market risk premium \( \gamma_{mt} \) by combining multiple quantiles of estimates

\[ \hat{\gamma}_{mt} = \sum_{j=1}^{n} \omega_j \gamma_{mt}(\tau_j) \quad \text{for } \tau_j = (0.1, 0.2, \cdots, 0.9) \]

3. We therefore compute the time series average of the cross sectional estimates

\[ \bar{\gamma}_m = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{mt} \]
4. The variance of $\hat{\gamma}_m$ can be constructed as

$$Var(\hat{\gamma}_m) = \frac{1}{T^2} \sum_{t=1}^{T} (\hat{\gamma}_{mt} - \bar{\gamma}_m)^2$$

The Adjust t-statistics can be obtained by applying Newey West(1987) statistics. Therefore, the market risk premium $\gamma_m$ can be estimated more efficiently in Fama MacBeth procedure. Our quantile combination approach can be extended to multifactor model where size and book to market factors can be included in the cross sectional regression.

Table 7 compare the robust estimation of Fama MecBeth cross section regression with OLS estimation. It is quite obvious that all weighted average quantile regression estimates of coefficients are lower than OLS estimator, which seems to indicate that the OLS estimation overestimate the true value of market risk premium under fat tail distribution of cross sectional stock return. Moreover, the size factor effect disappear in robust estimation. In OLS estimation, the size factor has strong significant (at 1% level) negative effect in explaining cross sectional return. In contrast, the robust estimator of size premium are insignificant in quantile combination estimation, which is consistent with the finding in Knez and Ready(1997) who found that the size effect almost disappear after excluding the most extreme observation.
6. Concluding Remark

The return distribution of security are significantly non-normal with thick tail and skewness. In the small sample, ordinary least square (OLS) estimator of beta risk would result in dramatic loss in efficiency when the data follow fat tail non-normal distribution. In this paper, we propose a novel quantile-based approach to estimate beta risk more efficiently without assuming a parametric structure on distribution. Multiple quantiles estimates are combined in weighting schemes to utilize information across the whole distribution. Our Monte Carlo Simulation strongly suggest that the proposal estimation strategies can substantially improve the efficiency of estimation (lower measurement error) for beta risk. This efficiency gain prevails even in the asymmetric distribution with positive (negative) skewness. Empirical studies also found a significant difference between our quantile-based estimator and OLS estimator for beta risk of individual firms, which may implies that the traditional OLS estimator may overestimate (underestimate) the true value of beta risk in the heavy tail distribution. Our robust approach also applied to Fama MacBeth second pass cross sectional test. The goodness of fit for the Fama MecBeth second pass cross sectional regression increase slightly as we use quantile combination approach instead of OLS to estimate beta risk, which implies a lower measurement error of beta in the first pass time series regression, leading to the improvement in the second pass cross section regression. Lastly, we studied the robust measurement of factor risk Premium and found that the size factor effect disappear after applying our robust approach to estimate risk premium. Since beta risk is latent variable which is unobservable, It is challenging to tell the reliability of beta estimators. Evaluating the performance of beta risk based on Asset pricing model can be misleading and spurious if the validity of CAPM model is not guaranteed. Anderson, Bollerslev and Diebold (2005,2003), Barndorff-Nielsen and Shephard (2004) demonstrate that the
realized beta estimated from high frequency return is consistent estimators of the true underlying integrated beta. Mathijs Cosemans et al (2010) therefore, construct an Out-of-Sample Beta Forecast by estimating the realized beta as proxy for true beta, which provide potential avenue for our further study.
Reference


Zhijie Xiao and Chi Wan 2011 "Robust Estimation of Conditional Volatility - A Quantile Regression Approach"