



Refining Our Understanding of Beta through Quantile Regressions

Working Paper Series—10-07 | April 2010

Allen B. Atkins

Northern Arizona University
The W. A. Franke College of Business

and

Pin Ng

Northern Arizona University
The W. A. Franke College of Business
20 W. McConnell Drive, Room 311
PO Box 15066
Flagstaff, AZ 86011-5066
Pin.Ng@nau.edu
928-523-8726
(Fax) 928-523-7331

Refining Our Understanding of Beta through Quantile Regressions

I. Introduction

Before Markowitz (1952, and others) investors concentrated on investment returns but did not think carefully about risk. Markowitz realized that risk was just as important of a factor when forming a portfolio and used variance of the return as a first risk measure. They measured past variance of a stock and assumed it would continue into the future. They also found that when stocks are combined into a portfolio, risk can be dramatically reduced. Bad news on one stock may be offset by good news on another stock, reducing variance. After combining about 15-20 stocks into a portfolio the benefits of diversification are mostly exhausted. The investor can no longer decrease risk simply by adding additional stocks. Essentially, the investor is left with market risk, also called non-diversifiable risk. Variance as a measure of risk is no longer adequate because it includes diversifiable risk. A new theory and measure of risk were needed.

In the field of finance, the Capital Asset Pricing Model (CAPM) was developed and has been a key theory since the 1960's (see Sharpe 1964 and others). One of its main contributions is to attempt to identify how the risk of a particular stock is related to the risk of the overall stock market. A main measure of this type of risk is called "beta". To estimate "beta", the returns of a particular stock (R_{it}) are regressed on the returns of a stock market proxy, say the S&P 500 (R_{mt}):

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}$$

The estimate of β_i is usually obtained using ordinary least squares (OLS) regression. The OLS estimates the responsiveness of R_{it} to changes in R_{mt} near the "center" of the R_{mt} distribution and provides an estimate of the average risk for the stock in relation to the risk of the overall market. If the estimated β_i is greater than (less than) 1.0, the stock is considered to be riskier (less risky) than the overall market on average and if the β_i is close to 1.0, the stock is about equally risky as the market on average.

If the relationship between a stock and the overall market exhibits heteroskedasticity, then more can be learned about the stock's behavior and its risk characteristics. Heteroskedasticity occurs when the variance of the error term ε_{it} varies across different values of R_{mt} . When the variance of the error term increases (decreases) as the overall market return increases, we call this "diverging (converging) heteroskedasticity". The quantile regression estimates of the beta slope under diverging (converging)

heteroskedasticity will generally increase (decrease) as we move from the lower towards the upper tail of the R_{it} distribution. These patterns of heteroskedasticity may provide information useful to investors. Investors may be able to select stocks that exhibit patterns that more closely match their preferences.

Kahneman and Tversky (1979) developed prospect theory which has been used to examine the behavioral aspects of stock market investors. One prediction that is made from their theory is that

- i) People exhibit “loss–aversion” in a gain frame, and
- ii) People exhibit “risk-seeking” in a loss frame.

In essence, if a person had done well or “gained”, they tend to be concerned about avoiding losses. Alternatively, if a person has experienced a bad outcome or has “lost”, their behavior may be risk seeking, perhaps trying to get back some of their losses. This implies that investors may prefer riskier stocks (higher beta) in down markets and less risk (lower beta) in up markets.

If a measure can be developed for each stock that incorporates investor’s preferences across good times and bad it could improve our ability to select portfolios that match investor’s preferences. When considering the beta of a stock, a flatter slope (safer) during good times and steeper slope (more risky) during bad times would match the Kahneman and Tversky findings. This would lead investors to prefer stocks that exhibit converging rather than diverging heteroscedasticity

In addition, investors may prefer a larger variance of return on their stocks when the market return is down and a smaller variance of return on their stocks when the market is up. This also would lead investors to prefer converging rather than diverging heteroskedasticity.

In the following section the data selection process is described. Then a discussion of quantile regressions is presented followed by a discussion of the results and their implications.

II. Data Selection

Stock prices are downloaded from the web site of Yahoo Finance. The data collected are from the beginning of January 1999 to the end of December of 2003 which represents 260 weeks of data. The total number of stocks initially in the sample was 6826. Of these 2956 were listed on Nasdaq and 3261 were listed on the NYSE and 609 were listed on the AMEX. Stocks were omitted from the sample if they had a price less than one dollar, or if they were not common shares. This left 3363 stocks that remained in the sample.

III. Quantile Regression

The quantile regression technique invented by Koenker and Bassett (1978) is a very powerful tool in uncovering heteroskedasticity in a regression model. In the CAPM model, given n observations of the individual return R_{it} and the market return R_{mt} for $t = 1, \dots, n$, the τ -th quantile regression coefficients, α_τ and β_τ , minimize the following objective function:

$$\sum_{t:R_{it}-\alpha_\tau-\beta_\tau R_{mt}\geq 0} \tau |R_{it}-\alpha_\tau-\beta_\tau R_{mt}| + \sum_{t:R_{it}-\alpha_\tau-\beta_\tau R_{mt}< 0} (1-\tau) |R_{it}-\alpha_\tau-\beta_\tau R_{mt}|$$

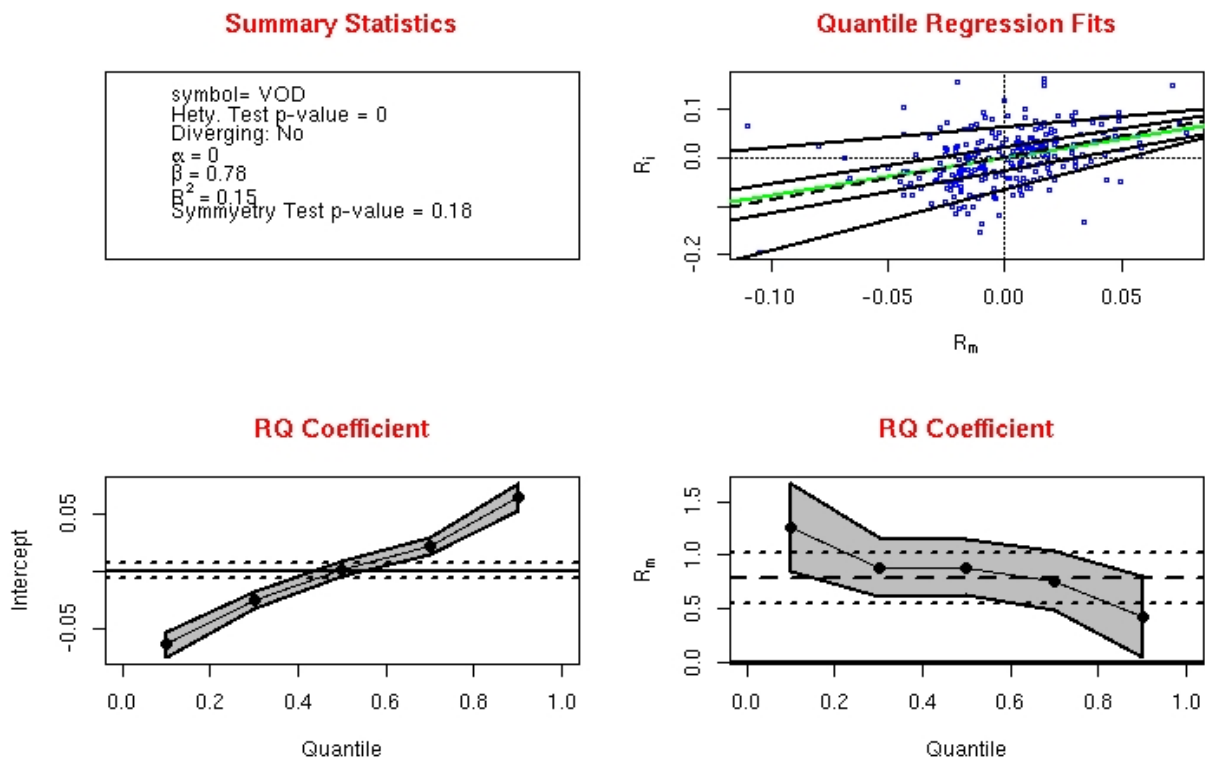
where $0 < \tau < 1$ determines the desired conditional quantile of interest. In the objective function, the positive and negative residuals, $R_{it} - \alpha_\tau - \beta_\tau R_{mt}$, receive different weights in the minimization process. All the positive residuals are assigned a weight of τ while the negative ones receive a weight of $(\tau - 1)$. Hence, $100\tau\%$ of the individual returns will fall above the τ -th quantile regression line $\alpha_\tau + x_i\beta_\tau$ and $100(1-\tau)\%$ below. Hence, the τ -th quantile regression line bisects the individual returns into two different portions, $100\tau\%$ and $100(1-\tau)\%$, conditioned on the various market returns. The special case of the 0.5-th quantile regression line, which is also the median regression line, divides the individual returns into two equal halves conditioned on the market returns so that half of the individual returns are above the line over the range of the market returns while the remaining half are below. The 0.1-th quantile regression line, on the other hand, divides the data such that only 10% of the individual returns fall below the line and 90% above while the 0.9-th quantile regression line will have 10% of the individual returns above and 90% below the line.

The upper-right panel of Figure 1 shows five different quantile regression lines for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 , and the different portions of the individual returns that fall above and below the lines are apparent. Also shown in green is the OLS regression, which provides the traditional estimate of the beta coefficient. The black dots connected by the solid line in the lower-right panel represent the regression quantile estimates of the beta coefficient β_τ for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . As we move from the left to the right with τ increases from 0.1 to 0.9, we can see that the quantile regression estimates of beta decrease from about $\beta_{0.1} = 1.25$ to $\beta_{0.9} = 0.35$ which reflects the declining slope of the quantile regression lines in the upper-right panel as we move from the lower quantile regression lines to the higher quantile regression lines. The grey band around the dots is the 95% confident band so that a particular τ -th quantile regression beta estimate β_τ is statistically different from 0 at a 5% level of significance when the

band does not contain 0 for the chosen τ . The horizontal dash line shows the value of the OLS estimated beta β_{OLS} with the horizontal dotted lines represent the 95% confidence band. The lower-left panel shows the regression quantile estimates of the alpha coefficient α_τ .

Figure 1. Stock with a strong converging heteroskedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r-square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.



IV. Heteroskedasticity

The presence of heteroskedasticity is apparent in Figure 1 from the increasing quantile regression beta coefficients β_τ . As the value of the market return on the horizontal axis increases, the degree of variation of the individual return decreases. We call this sort of heteroskedasticity a “converging heteroskedasticity” while the form of heteroskedasticity depicted in Figure 2 is termed “diverging heteroskedasticity”. The strength of the heteroskedasticity is classified using the p -value for the Wald test for existence of heteroskedasticity introduced by Koenker and Bassett (1982). Since the smaller the p -value, the stronger is the evidence in the data against the null hypothesis of the absence of heteroskedasticity, we classify the strength of heteroskedasticity using the delineation specified in Table 1.

Figure 3 and 4 provide examples of sample stock returns with moderate and weak converging heteroskedasticity according to the increasingly larger p -value shown among the summary statistics in the upper-left panel of the figures while Figure 5 and 6 show examples of moderate and weak converging heteroskedasticity.

We use the **rq** function in the **quantreg** package (Koenker, 2010) available from the GNU Free Software **R** for statistical computing and graphics (R Development Core Team, 2008) to compute the quantile regression coefficients in this study. The function uses a modified version of Barrodale and Roberts's (1974) algorithm for L_1 regression as described in Koenker and d'Orey (1987, 1994) for small sample sizes and the interior-point algorithm described in Koenker and Ng (2005) for large sample sizes. The standard error assumes local linearity of the conditional quantile functions and computes an Eicker-Huber-White sandwich estimate using a local estimate of the sparsity as described in Koenker (2005). Koenker and Hallock (2001) is an excellent non-technical primer for quantile regression.

Figure 2. Strong diverging hetorskedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r-square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.

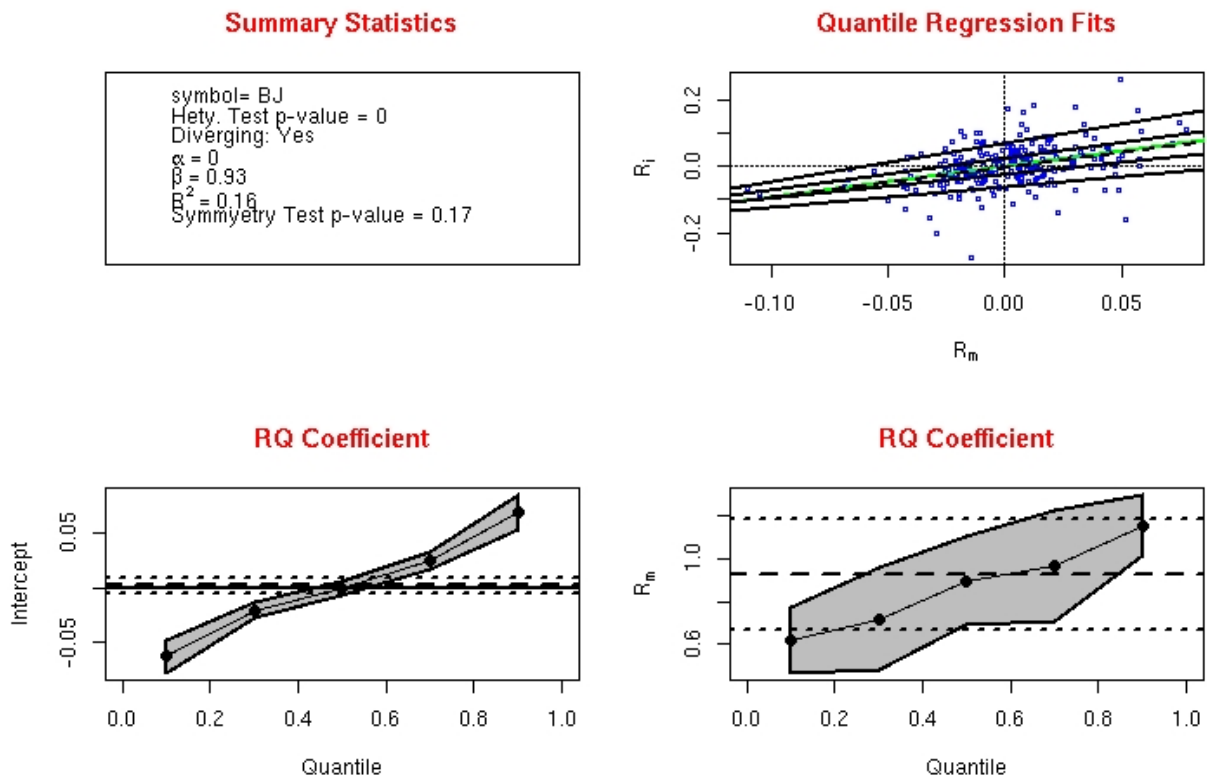


Table 1. Classification of heteroskedasticity.

The table classifies the degree of converging and diverging heteroskedasticity into strongly, moderately, weakly and little or no heteroskedasticity. The smaller the p -value is for the Wald test for heteroskedasticity, the stronger is the degree of diverging or converging. Divergence or convergence is determined by the difference between the higher ($\beta_{0.9}$) and lower regression quantiles ($\beta_{0.1}$) for beta. A negative difference with $\beta_{0.9} - \beta_{0.1} < 0$ is defined as converging while a positive difference $\beta_{0.9} - \beta_{0.1} > 0$ is called diverging.

P value	Type of Heteroskedasticity	$\beta_{0.9} - \beta_{0.1}$	% of sample	Number of stocks
[0.0-0.1]	Strongly Converging	-	2.9	97
[0.1-0.2]	Moderately Converging	-	3.1	104
[0.2-0.3]	Weakly Converging	-	3.8	127
[0.3-1.0]	Little or no heteroskedasticity	- or +	70.4	2368
[0.2-0.3]	Weakly Diverging	+	6.8	228
[0.1-0.2]	Moderately Diverging	+	6.2	208
[0.0-0.1]	Strongly Diverging	+	6.8	228

Figure 3. Moderate diverging heteroskedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r -square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.

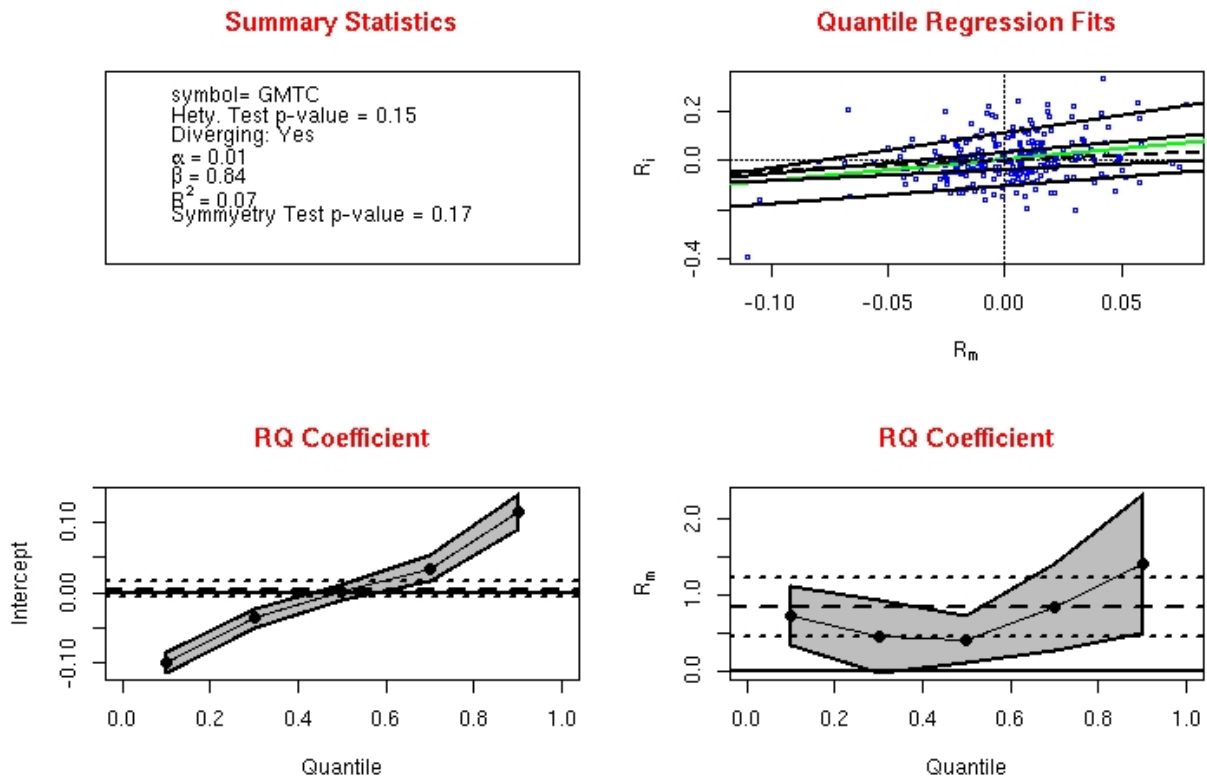


Figure 4. Weak diverging heterkedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r -square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.

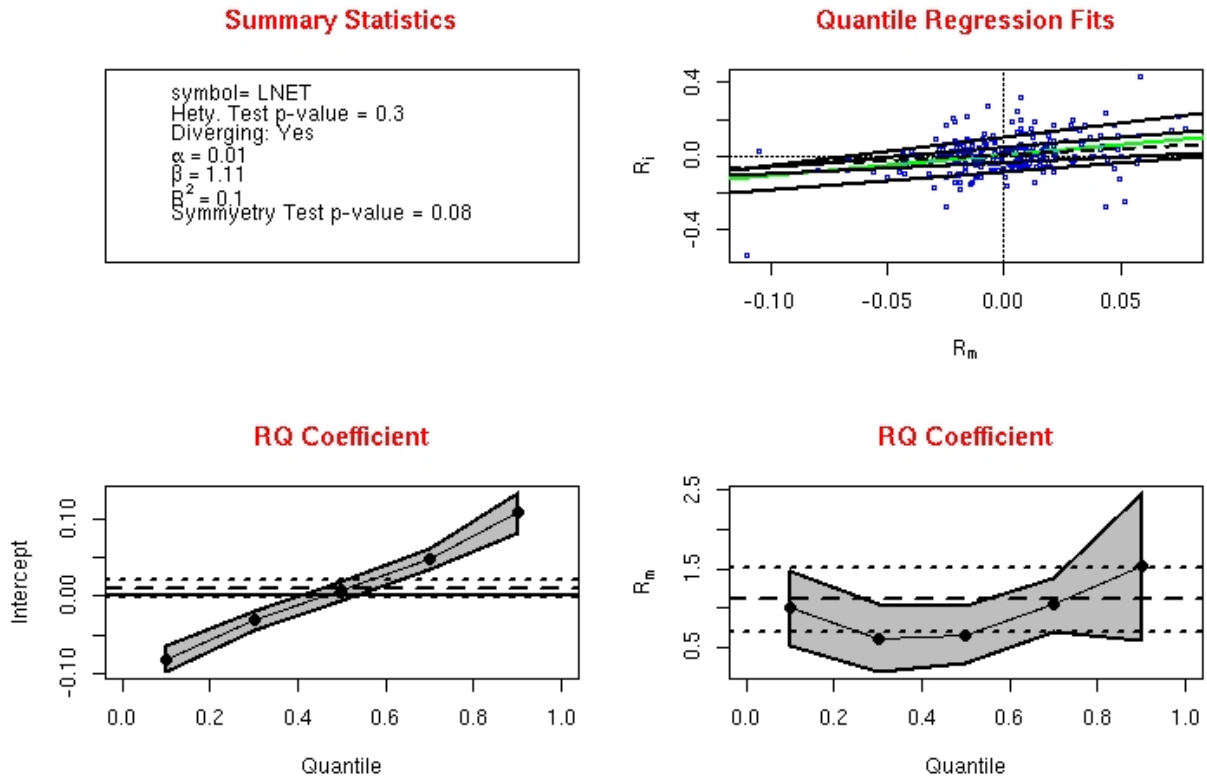


Figure 5. Moderate converging heteroskedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r -square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.

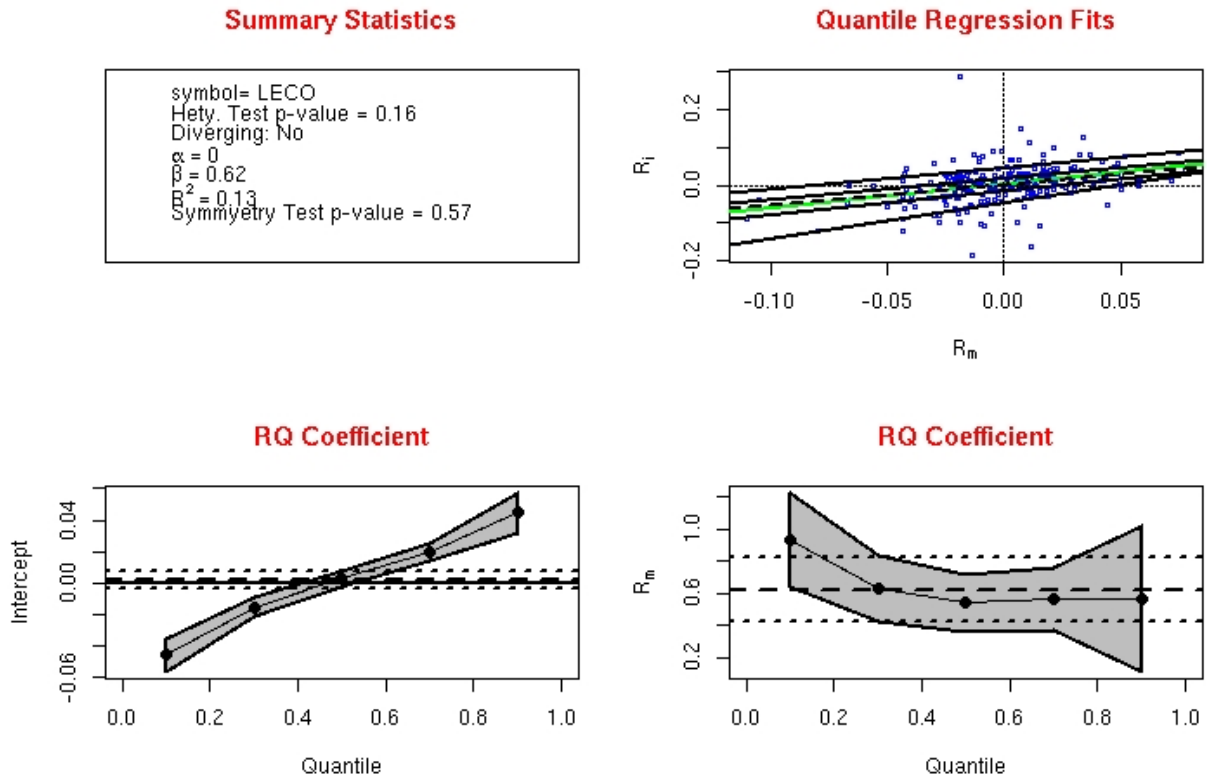
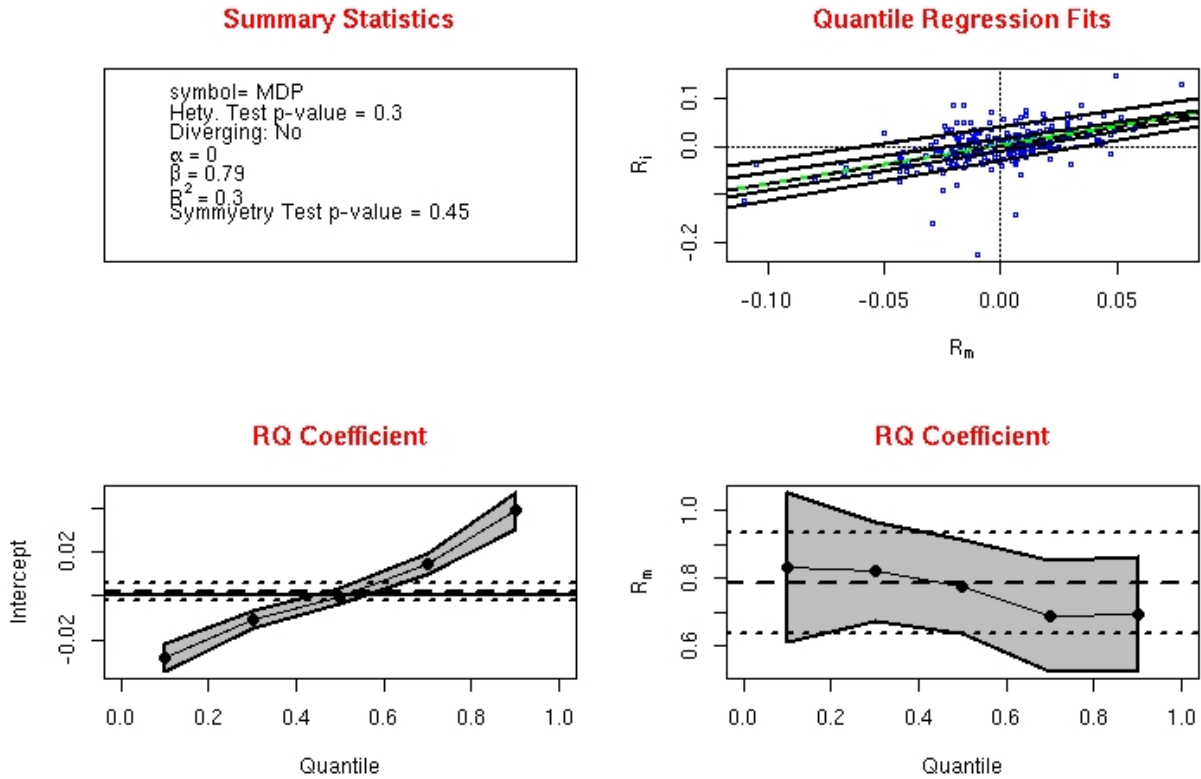


Figure 6. Weak converging heteroskedasticity.

The upper-left panel contains summary statistics of the chosen stock, which include the ticker symbol, p -value for the Wald test for heteroskedasticity, indicator for whether the stock is diverging, estimated Alpha and beta, r -square, and the p -value for the test for symmetry of the stock return. The upper-right panel shows the various quantile regression fits of the CAPM model for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 . The lower-left panel contains the quantile regression estimates of Alpha while the lower-right panel shows the quantile regression estimates of beta for $\tau = 0.1, 0.3, 0.5, 0.7$ and 0.9 along with their 95% confidence band. The dash line indicates the magnitude of the OLS beta estimate while the dotted lines are the 95% confidence interval.



V. Interpretation of the Results

The sample of 3363 stocks is tested for levels of converging or diverging heteroskedasticity. The Wald test for the existence of heteroskedasticity is used to determine the type and level of heteroskedasticity and the results are presented in Table 1. Strong, moderate and weak heteroskedasticity are defined as p -values between 0.0-0.1, 0.1-0.2 and 0.2-0.3, respectively. For the sample of 3363 stocks 2.9%, 3.1% and 3.8% showed strongly converging, moderately converging and weakly converging heteroskedasticity, respectively. In total, 9.8% of the sample (328 stocks) showed some degree of converging heteroskedasticity. Conversely, 6.8%, 6.2% and 6.8% showed strongly, moderately and weakly diverging heteroskedasticity. In total, 19.8% of the sample (664 stocks) showed some degree of diverging heteroskedasticity. The remaining 70.4% of the sample showed no significant heteroskedasticity.

One interpretation of these findings is that just under 10% of the stocks will fit the behavioral preference indicated by the Kahneman and Tversky findings. These stocks, showing converging heteroskedasticity, would be more preferred by investors than the other 90% of the stocks if all other

factors are held constant. On the other side, just under 20% of the stocks show patterns that are contrary to the Kahneman and Tversky findings. These stocks, showing diverging heteroskedasticity, would be less preferred by investors if all other factors are held constant.

Investors are always looking for any possible indicator that can help them improve their portfolios. Even small improvements can provide benefits to investors' utility. The results presented may give new information about 30% of the universe of stocks that investors will choose from. When a portfolio can be chosen that more aligns with investors preferences, it may be a benefit to investors.

References

- Barroda , I. and F. D. K. Roberts (1974), “Solution of an overdetermined system of equations in the l1 norm”, *Communications of the ACM*, Vol.17 No.6, pp.319-320.
- Kahneman, Daniel, and Amos Tversky (1979), “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica*, Vol. 47, No. 2, pp. 263-291.
- Koenker, Roger. (2005), *Quantile Regression*, Cambridge University Press, New York.
- . (2009), *quantreg: Quantile Regression*, R package version 4.26, <http://www.r-project.org>.
- and Gilbert Bassett (1978), “Regression Quantiles”, *Econometrica*, Vol. 46, No. 1, pp. 33-50.
- and Gilbert Bassett (1982), “Robust Tests for Heteroscedasticity Based on Regression Quantiles, *Econometrica*”, Vol. 50, No. 1, pp. 43-61.
- and V. d’Orey (1987), “Algorithm AS 229: Computing Regression Quantiles”, *Applied Statistics*, Vol. 36, No.3, pp. 383-393.
- and V. d’Orey (1994), “Remark AS R92: A Remark on Algorithm AS 229: Computing Dual Regression Quantiles and Regression Rank Scores”, *Applied Statistics*, Vol. 43, No.2, 410-414.
- , K. Hallock, (2001), “Quantile regression: an introduction.” *Journal of Economic Perspectives*, Vol. 15, pp. 143-156.
- and P. Ng (2005), “Frisch-Newton Algorithm for Sparse Quantile Regression, *Acta Mathematicae Applicatae Sinica (English series)*, Vol. 21, 225-236.
- Markowitz, Harry (1952), “Portfolio Selection”, *The Journal of Finance*, Vol. 7, No. 1, pp. 77-91.
- Sharpe, William F. (1964), “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk”, *The Journal of Finance*, Vol. 19, No. 3, pp. 425-442.