Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regression[*]

Tae-Hwan Kim
School of Economics, University of Nottingham
University Park, Nottingham NG7 2RD, UK
(Phone) 44-115-951-5466   (Fax) 44-115-951-4159
Tae-Hwan.Kim@nottingham.ac.uk

Halbert White
Department of Economics, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0508
(Phone) 619-534-3502   (Fax) 619-534-7040
hwhite@weber.ucsd.edu

First Draft: January 2000

This Version: April 2002

Abstract: To date the literature on quantile regression and least absolute deviation regression has assumed either explicitly or implicitly that the conditional quantile regression model is correctly specified. When the model is misspecified, confidence intervals and hypothesis tests based on the conventional covariance matrix are invalid. Although misspecification is a generic phenomenon and correct specification is rare in reality, there has to date been no theory proposed for inference when a conditional quantile model may be misspecified. In this paper, we allow for possible misspecification of a linear conditional quantile regression model. We obtain consistency of the quantile estimator for certain "pseudo-true" parameter values and asymptotic normality of the quantile estimator when the model is misspecified. In this case, the asymptotic covariance matrix has a novel form, not seen in earlier work, and we provide a consistent estimator of the asymptotic covariance matrix. We also propose a quick and simple test for conditional quantile misspecification based on the quantile residuals.

Key Words: Conditional Quantile, Misspecification, Asymptotic Normality, Asymptotic Covariance Matrix.

---

1. Introduction

Since the seminal work of Koenker and Bassett (1978) and Bassett and Koenker (1978), the literature on quantile regression and least absolute deviation (LAD) regression has grown rapidly in many interesting directions, such as simultaneous equation and two stage estimation [Amemiya (1982), Powell (1983)], censored regression [Powell (1984), Powell (1986), Buchinsky and Hahn (1998)], serial correlation and GLS estimation [Weiss (1990)], bootstrap methods [Hahn (1995), Horowitz (1998)], structural break testing [Bai (1995)], ARCH models [Koenker and Zhao (1996)], and unit root testing [Herce (1996)].

All these papers, however, assume explicitly or implicitly that the conditional quantile regression model is correctly specified. When the model is misspecified, confidence intervals and hypothesis tests based on the conventional covariance matrix are, as we show, invalid. Even though misspecification is a generic phenomenon and correct specification is rare in reality, there has to date been no theory proposed for inference when a conditional quantile model may be misspecified. In this paper, we allow for possible misspecification of a linear conditional quantile regression model. We obtain consistency of the quantile estimator for certain "pseudo-true" parameter values and asymptotic normality of the quantile estimator when the model is misspecified. In this case, the asymptotic covariance matrix has a novel form, not seen in earlier work, and we provide a consistent estimator of the asymptotic covariance matrix.

Of course, one can estimate the conditional quantile model without assuming correct specification using various non-parametric methods such as kernel estimation [Sheather and Marron (1990)], nearest-neighbor estimation [Bhattacharya and Gangopadhyay (1990)], or using artificial neural networks [White (1992)]. Our results thus provide a convenient parametric alternative to nonparametric methods when researchers are not sure about correct specification or when they want to keep a parametric model for reasons of parsimony or interpretability even though it may not pass a specification test such as the nonparametric kernel based test proposed by Zheng (1998).

2. Basic Assumptions and Model

Consider a random series (Y_t, X_t') where $t=1,2,...,T$, Y_t is a scalar, X_t is a $k \times 1$ vector, and $k \equiv \tilde{k} + 1$. The first element in X_t is one for all t. First, we specify the data generating process.

Assumption 1. *The sequence $\{Y_t, X_t'\}$ is independent and identically distributed (iid).*

The iid assumption is made for clarity and simplicity. It can be straightforwardly relaxed. We denote the conditional distribution of Y_t given $X_t = x$ by $F_{Y|X}(\cdot \mid x)$. As is now standard in the quantile regression literature, we define the "check" function $r_q : R \to R^+$ for given $q \in (0,1)$ as

$$r_q(z) \equiv z y_q(z),$$

where $y_q(z) \equiv q - 1_{[z \leq 0]}$. We define the qth conditional quantile of Y_t given X_t as

$$q_q(Y_t \mid X_t) \equiv \inf\{ y : F_{Y|X}(y \mid X_t) \geq q\}.$$

Next, we impose the following assumption on the joint density of (Y_t, X_t').

Assumption 2. *The random vector $(Y_t, X_{t2}, X_{t3},..., X_{tk})$ is continuously distributed with joint probability density $f_{Y,X}(\cdot,\cdot)$ and conditional probability density $f_{Y|X}(\cdot \mid x)$ for Y_t given $X_t = x$.*

Under Assumption 2, the conditional quantile $q_q(Y_t \mid X_t)$ satisfies $\int_{-\infty}^{q_q(Y_t|X_t)} f_{Y|X}(y \mid X_t)dy - q = 0$, which is equivalent to

$$E[y_q(Y_t - q_q(Y_t \mid X_t)) \mid X_t] = 0. \tag{2.1}$$

The condition (2.1) can be used to check if a given function of X_t is the qth conditional quantile of Y_t given X_t or not.

It is well known that

$$q_q(Y_t \mid X_t) \in \underset{f}{\arg\min} \; E[r_q(Y_t - f(X_t))] \tag{2.2}$$

3

where f belongs to a space of measurable functions defined as $F \equiv \{g : R^k \to R \text{ such that } g \text{ is measurable and } E \,|\, g(X_t)\,| < \infty \}$. Here, we focus only on the affine space A ($\subset F$) $\equiv \{g : R^k \to R \text{ such that for some } k \times 1 \text{ vector } \boldsymbol{b}, \ g(X_t) = X_t'\boldsymbol{b} \text{ and } E \,|\, g(X_t)\,| < \infty \}$. For the objective function in (2.2) to be well-defined, we impose some moments conditions on (Y_t, X_t'). The following conditions are sufficient.

Assumption 3. $E \,|\, Y_t \,| < \infty$, and $E \,\|\, X_t \,\| < \infty$, where $\|\, X_t \,\| \equiv (X_t'X)^{1/2}$.

Now we give a definition of correct model specification.

Definition 1. *We say a conditional quantile model $\{h(\cdot, \boldsymbol{b}) \in F, \ \boldsymbol{b} \in R^k \}$ is correctly specified for $q_{\boldsymbol{q}}(Y_t \,|\, X_t)$ if and only if there exists a vector $\boldsymbol{b}_0 \in R^k$ such that $h(X_t, \boldsymbol{b}_0) = q_{\boldsymbol{q}}(Y_t \,|\, X_t)$ almost surely, i.e. $q_{\boldsymbol{q}} \in F$.*

We impose the following quantile version of the orthogonality condition.

Assumption 4. *There exists $\boldsymbol{b}^*$ such that $E(\boldsymbol{y}_{\boldsymbol{q}}(Y_t - X_t'\boldsymbol{b}^*)X_t) = 0$.*

Given the "pseudo-true parameters" $\boldsymbol{b}^*$ of Assumption 4, we can define the "error" $e_t \equiv Y_t - X_t'\boldsymbol{b}^*$. Assuming that $E(\boldsymbol{y}_{\boldsymbol{q}}(e_t) \,|\, X_t) = 0$, which is stronger than Assumption 4, is equivalent to assuming that the conditional quantile model is correctly specified. This can be easily checked using (2.1). Thus, Assumption 4 permits the conditional quantile model to be misspecified. Let B denote a subset in R^k large enough to contain $\boldsymbol{b}^*$. Then, under our conditions Assumption 4 implicitly defines the parameter of interest $\boldsymbol{b}^*$ as the solution to the minimization problem:

$$\min_{\boldsymbol{b} \in B} E[\boldsymbol{r}_{\boldsymbol{q}}(Y_t - X_t\boldsymbol{b})],$$

which is well-defined by Assumption 3. As discussed by White (1994, pp. 74-75), this optimization problem corresponds to maximiming the expected log-likelihood for a particular

4

density function. The regression quantile parameter estimator $\hat{\boldsymbol{b}}_T$, obtained by minimizing the sample analog

$$\min_{\boldsymbol{b} \in B} S_T(\boldsymbol{b}) \equiv T^{-1} \sum_{t=1}^{T} \boldsymbol{r_q}(Y_t - X_t \boldsymbol{b}), \tag{2.3}$$

can therefore be viewed as a quasi-maximum likelihood estimator (QMLE).

3. Consistency and Asymptotic Normality

First, we establish the consistency of the quantile estimator $\hat{\boldsymbol{b}}_T$ for $\boldsymbol{b}^*$ using Lemmas 2.2 and 2.3 in White (1980a). The consistency result is the first step in deriving the asymptotic normality of the quantile estimator. The following additional assumptions suffice for the proof of consistency.

Assumption $3'$**.** *There exists* $\boldsymbol{d} > 1$ *such that* $E(\| X_t \|^{\boldsymbol{d}}) < \infty$.

Assumption 5. $\boldsymbol{b}^* \in B$, *where* B *is a compact subset of* R^k.

Assumption 6. *For all* x, $f_{\boldsymbol{e}|X}(0 \,|\, x) > 0$ *where* $f_{\boldsymbol{e}|X}(\cdot \,|\, x)$ *is the conditional density of* $\boldsymbol{e}_t \equiv Y_t - X_t'\boldsymbol{b}^*$ *given* $X_t = x$.

Lemma 1. *Suppose that Assumptions 1, 2, $3'$, 4, 5, and 6 hold. Then* $\hat{\boldsymbol{b}}_T - \boldsymbol{b}^* = o_p(1)$.

All proofs are provided in the Mathematical Appendix.

There are several techniques available in the literature to derive asymptotic normality for the quantile and the LAD estimators. Among these are the linear programming method [Koenker and Bassett (1978), Bassett and Koenker (1978)]; the smoothing method [Bloomfield and Steiger (1983)]; the convexity method [Pollard (1991)]; and the generalized Taylor expansion method [Phillips (1991)]. Here, we follow the method used by Huber (1967), Ruppert and Carroll (1980), and later extended by Pollard (1985). Huber (1967) gave sufficient conditions that deliver asymptotic normality for any sequence $\hat{\boldsymbol{b}}_T$ satisfying

$$T^{-1/2} \sum_{t=1}^{T} \boldsymbol{j}(Z_t, \hat{\boldsymbol{b}}_T) = o_p(1).$$ (3.1)

In our case, $\boldsymbol{j}(Z_t, \hat{\boldsymbol{b}}_T) = X_t \boldsymbol{y}_q (Y_t - X_t' \hat{\boldsymbol{b}}_T)$. The condition (3.1) can be viewed as the first order condition for the quantile estimator $\hat{\boldsymbol{b}}_T$ because the left hand term is essentially the vector of left partial derivatives of the objective function in (2.3) evaluated at $\hat{\boldsymbol{b}}_T$. Because we define the quantile estimator $\hat{\boldsymbol{b}}_T$ using (2.3) rather than (3.1), we must establish (3.1). For this, we use the following assumption.

Assumption 3″. *There exists $\boldsymbol{d} > 2$ such that $E(\| X_t \|^{\boldsymbol{d}}) < \infty$.*

Lemma 2. *Suppose that Assumptions 2 and 3″ hold. Then,*

$$T^{-1/2} \sum_{t=1}^{T} X_t \boldsymbol{y}_q (Y_t - X_t' \hat{\boldsymbol{b}}_T) = o_p(1).$$

Next, we define $\boldsymbol{l}(\boldsymbol{b}) \equiv E[X_t \boldsymbol{y}_q (Y_t - X_t' \boldsymbol{b})]$. Since $\boldsymbol{l}(\boldsymbol{b})$ can be shown to be continuously differentiable in $\boldsymbol{b}$, we have by the mean value theorem

$$\boldsymbol{l}(\boldsymbol{b}) = \boldsymbol{l}(\boldsymbol{b}^*) - Q^*(\boldsymbol{b} - \boldsymbol{b}^*),$$ (3.2)

where $Q^* \equiv E[f_{e|X}(\boldsymbol{l}_t^* \mid X_t) X_t X_t']$ and $\boldsymbol{l}_t^*$ is between 0 and $X_t'(\boldsymbol{b} - \boldsymbol{b}^*)$. Note that $\boldsymbol{l}(\boldsymbol{b}^*) = 0$ by the definition of $\boldsymbol{b}^*$. We use the expression for $\boldsymbol{l}(\boldsymbol{b})$ in (3.2) to obtain a variant of the Taylor expansion of $\boldsymbol{l}(\boldsymbol{b})$, which will be the key step obtaining the asymptotic distribution of the quantile estimator. We impose the following conditions.

Assumption 3‴. $E(\| X_t \|^3) < \infty$.

Assumption 7. *The conditional density $f_{e|X}(\boldsymbol{l} \mid x)$ of e_t given $X_t = x$ is Lipschitz continuous: i.e. $| f_{e|X}(\boldsymbol{l}_1 \mid x) - f_{e|X}(\boldsymbol{l}_2 \mid x) | \leq L_0 | \boldsymbol{l}_1 - \boldsymbol{l}_2 |$ for some constant $0 < L_0 < \infty$ and for all x.*

Lemma 3. *Suppose that Assumption 3‴ and 7 hold. Then*

$$\| \boldsymbol{l}(\boldsymbol{b}) - \boldsymbol{l}(\boldsymbol{b}^*) + Q_0(\boldsymbol{b} - \boldsymbol{b}^*) \| = o(\| \boldsymbol{b} - \boldsymbol{b}^* \|),$$

where $Q_0 \equiv E[f_{e|X}(0 \mid X_t) X_t X_t']$.

The final step in obtaining asymptotic normality of the quantile estimator $\hat{\boldsymbol{b}}_T$ is to show that $T^{1/2} \boldsymbol{l}(\hat{\boldsymbol{b}}_T)$ converges to some random variable in distribution. In fact, using Theorem 3 in Huber (1967), $-T^{1/2} \boldsymbol{l}(\hat{\boldsymbol{b}}_T)$ and $T^{-1/2} \sum_{t=1}^{T} X_t \boldsymbol{y}_q(\boldsymbol{e}_t)$ turn out to be asymptotically equivalent. To show this, we impose some additional assumptions.

Assumption 8. $\boldsymbol{b}^*$ *is an interior point of* B.

Assumption 9. Q_0 *is positive definite.*

Assumption 10. *There exists a constant* f_1 *such that* $f_{e|X}(\boldsymbol{l} \mid x) \le f_1$ *for all* $\boldsymbol{l}$ *and* x.

Lemma 4. *Suppose that Assumptions 1, 2, 3‴, 4-6 and 8-10 hold. Then*

$$T^{-1/2} \sum_{t=1}^{T} X_t \boldsymbol{y}_q(\boldsymbol{e}_t) + T^{1/2} \boldsymbol{l}(\hat{\boldsymbol{b}}_T) = o_p(1).$$

Combining Lemma 3 and Lemma 4 and adding one more assumption permit us to state our main theorem. We impose

Assumption 11. $V \equiv E(\boldsymbol{y}_q(\boldsymbol{e}_t)^2 X_t X_t')$ *is positive definite.*

Theorem 1. *Suppose that Assumptions 1, 2, 3‴, 4-11 hold. Then,*

$$T^{1/2}(\hat{\boldsymbol{b}}_T - \boldsymbol{b}^*) \xrightarrow{d} N(0, C)$$

where $C \equiv Q_0^{-1} V Q_0^{-1}$.

The asymptotic distribution in Theorem 1 includes all previously obtained results as special cases. Suppose that the linear conditional quantile model is correctly specified. Then, it can be shown that $V = \mathbf{q}(1-\mathbf{q})Q$ where $Q \equiv E(X_t X_t')$. This case thus corresponds to Powell (1984) who obtains $T^{1/2}(\hat{\mathbf{b}}_T - \mathbf{b}^*) \overset{d}{\to} N(0, \mathbf{q}(1-\mathbf{q})Q_0^{-1}QQ_0^{-1})$. Consider the more restricted case where not only is the linear conditional quantile model correctly specified but also there is no conditional heterogeneity in the density f at the origin (no "heteroaltitudinality"), that is $f_{e|X}(0|X_t) = f(0)$. Then, one can show that $Q_0 = f(0)Q$ as well as $V = \mathbf{q}(1-\mathbf{q})Q$. This corresponds to the standard case of Koenker and Bassett (1978), who obtain

$$T^{1/2}(\hat{\mathbf{b}}_T - \mathbf{b}^*) \overset{d}{\to} N(0, \frac{\mathbf{q}(1-\mathbf{q})}{f(0)^2}Q^{-1}).$$

Now consider testing a hypothesis about $\mathbf{b}^*$:

$H_0 : R\mathbf{b}^* = r$

$H_a : R\mathbf{b}^* \neq r$,

where R is a finite $q \times k$ matrix of full row rank and r is a finite $q \times 1$ vector. Then Theorem 1 implies that, for example,

$$T(R\hat{\mathbf{b}}_T - r)'[R\hat{C}_T^{-1}R']^{-1}(R\hat{\mathbf{b}}_T - r) \overset{d}{\to} \mathbf{c}_q^2 \tag{3.3}$$

under the null hypothesis, where $\hat{C}_T \overset{p}{\to} C$ is a covariance matrix estimator, consistent for C despite the possible misspecification. To implement such tests we require a consistent estimator $\hat{C}_T$. This is the focus of our next section.

4. Consistent Covariance Matrix Estimation

In this section, we provide an estimator for the asymptotic covariance matrix C that is consistent despite possible misspecification. The asymptotic covariance matrix consists of two components: Q_0 and V. Powell (1984) suggested the following estimator for Q_0 without formally proving its consistency:

$$\hat{Q}_{0T} \equiv (2\hat{c}_T T)^{-1} \sum_{t=1}^{T} 1_{[-\hat{c}_T \le \hat{e}_t \le \hat{c}_T]} X_t X_t' \tag{4.1}$$

where $\hat{c}_T$ may be a function of the data and $\hat{e}_t \equiv Y_t - X_t'\hat{b}_T$. We impose the following conditions on the sequence $\{\hat{c}_T\}$.

Assumption 12. *There is a stochastic sequence $\{\hat{c}_T\}$ and a non-stochastic sequence $\{c_T\}$ such that*

(i) $\quad \dfrac{\hat{c}_T}{c_T} \xrightarrow{p} 1$

(ii) $\quad c_T = o(1)$

(iii) $\quad c_T^{-1} = o(\sqrt{T})$.

We can now rigorously establish the consistency of Powell's estimator $\hat{Q}_{0T}$.

Lemma 5. *Suppose that $T^{1/2} \| \hat{b}_T - b^* \| = O_p(1)$ and Assumptions 1, 3''', 7, 10 and 12 hold. Then $\hat{Q}_{0T} \xrightarrow{p} Q_0$.*

Next, we use the plug-in principle to propose a consistent estimator of V:

$$\hat{V}_T \equiv T^{-1} \sum_{t=1}^{T} \hat{u}_t^2 X_t X_t',$$

where $\hat{u}_t \equiv y_q(\hat{e}_t)$. The estimator $\hat{V}_T$ is completely analogous to White's (1980b) estimator, for which $\hat{u}_t$ is the OLS residual.

Lemma 6. *Suppose that $T^{1/2} \| \hat{b}_T - b^* \| = O_p(1)$ and Assumptions 1, 3''', and 10 hold. Then*

$$\hat{V}_T \xrightarrow{p} V.$$

We now define our estimator for the asymptotic covariance matrix C as

$$\hat{C}_T \equiv \hat{Q}_{0T}^{-1}\hat{V}_T\hat{Q}_{0T}^{-1}.\tag{4.2}$$

Together, Lemmas 5 and 6 imply $\hat{C}_T \xrightarrow{p} C$, ensuring the consistency of the covariance estimator. This allows us to obtain a computable Wald statistic and its asymptotic distribution as follows.

Theorem 2. *Suppose that Assumptions 1, 2, $3'''$, 4-12 hold. Let the null hypothesis be given by* $H_0 : R\boldsymbol{b}^* = r$ *where R is a given finite $q \times k$ matrix of full row rank and r is a given finite $q \times 1$ vector. Then, under H_0,*

$$T(R\hat{\boldsymbol{b}}_T - r)'[R\hat{C}_T^{-1}R']^{-1}(R\hat{\boldsymbol{b}}_T - r) \xrightarrow{d} \boldsymbol{c}_q^2,$$

where $\hat{C}_T$ is defined in (4.2).

Although we do not pursue the issue here, we note that just as MacKinnon and White (1985) found modifications of $\hat{V}_T$ that afforded improvements in finite sample properties, so also may there be analogous modifications of $\hat{V}_T$ here. We leave this to subsequent research.

5. A Test for Correct Quantile Specification

As we have seen in Section 3, if the conditional quantile model is correctly specified, then we have the quantile version of information matrix equality: $V = \boldsymbol{q}(1-\boldsymbol{q})Q$. We formally state this in the following lemma.

Lemma 7. *Suppose that Assumptions 2 and $3''$ hold and let $u_t \equiv \boldsymbol{y}_q(\boldsymbol{e}_t)$. If the conditional quantile model is correctly specified (i.e. $E(u_t \mid X_t) = 0$), then we have $V = \boldsymbol{q}(1-\boldsymbol{q})Q$: that is, $E(u_t^2 X_t X_t') = E(u_t^2)E(X_t X_t')$. Equivalently, this can be written $E[(u_t^2 - \boldsymbol{s}_u^2)X_t X_t'] = 0$, where $\boldsymbol{s}_u^2 \equiv E(u_t^2) = \boldsymbol{q}(1-\boldsymbol{q})$.*

It is interesting to note that in order to have the same equality as in Lemma 7 in the context of OLS regression, we require two conditions: (1) the conditional expectation model is correctly

10

specified and (2) that there is no conditional heteroskedasticity in u_t. See White (1980b) for details. However, for quantile regression, correct specification is the only sufficient condition due to the special structure of u_t. According to Lemma 7, any misspecification in the conditional quantile is a form of conditional heteroskedasticity in u_t. In such situations, inference based on the information equality is invalid.

Zheng (1998) has developed a consistent test for conditional quantile misspecification. That test is based on a nonparametric kernel estimation and may accordingly be somewhat cumbersome to implement. We now propose a very easy to use specification testing procedure that exploits the quantile version of the information matrix equality in Lemma 7. Under the null hypothesis of correct specification of the conditional quantile function, we have

$$E[h(X_t)(u_t^2 - \mathbf{s}_u^2)] = 0, \tag{5.1}$$

where $h(\cdot)$ is a measurable $s \times 1$ vector function. The information matrix test obtains when $h(X_t)$ selects some elements from the matrix $X_t X_t'$. One example is $h(X_t) = vech(JX_t X_t' J')$ with $J = [0_{\tilde{k} \times 1} \ I_{\tilde{k} \times \tilde{k}}]$. In this example $h(X_t)$ is a $\tilde{k}(\tilde{k}+1)/2 \times 1$ vector containing all the elements in the lower triangle and diagonal of the $\tilde{k} \times \tilde{k}$ matrix $JX_t X_t' J'$. Using the fact that $u_t^2 = \mathbf{s}_u^2 - (1 - 2\mathbf{q})\mathbf{y}_\mathbf{q}(\mathbf{e}_t)$, the expression in (5.1) can be shown to be equivalent to

$$E[h(X_t)\mathbf{y}_\mathbf{q}(\mathbf{e}_t)] = 0. \tag{5.2}$$

A scaled sample version of the expectation in (5.2), which is our proposed test statistic for the null of correct specification, is given by

$$T^{-1/2} m(\hat{\mathbf{b}}_T),$$

where $m(\hat{\mathbf{b}}_T) = T^{-1} \sum_{t=1}^{T} h(X_t)\mathbf{y}_\mathbf{q}(Y_t - X_t'\hat{\mathbf{b}}_T)$.

We restrict the space to which the selection function $h(\cdot)$ belongs using the following moment condition.

Assumption 13. *There exists $\mathbf{d} > 2$ such that $E(\| h(X_t) \|^\mathbf{d}) < \infty$.*

The following lemma is useful in deriving the asymptotic distribution of the proposed statistic.

Lemma 8. *Suppose that Assumptions 1, $3'''$ and 13 hold. Then,*

$$\sup_{\boldsymbol{b} \in B_T} |T^{1/2}[m(\boldsymbol{b}) - m(\boldsymbol{b}^*)] - T^{-1/2} \sum_{t=1}^{T} h(X_t)[F_t(\boldsymbol{b}) - F_t(\boldsymbol{b}^*)]| = o_p(1)$$

where $B_T \equiv \{\boldsymbol{b} \in B : T^{1/2} \| \boldsymbol{b} - \boldsymbol{b}^* \| \leq M_T$ *and* $M_T = O(1)\}$, *and* $F_t(\boldsymbol{b}) \equiv F_{Y|X}(X_t'\boldsymbol{b} \mid X_t)$.

The proof easily follows from Andrews (1989) once we show that

$$T^{1/2}[m_T(\boldsymbol{b}) - m_T(\boldsymbol{b}^*)] - T^{-1/2} \sum_{t=1}^{T} h(X_t)[F_t(\boldsymbol{b}) - F_t(\boldsymbol{b}^*)] = T^{-1/2} \sum_{t=1}^{T} H(Z_t, \boldsymbol{b})$$

where $H(Z_t, \boldsymbol{b}) \equiv h(X_t)\{1_{[Y_t \leq X_t'\boldsymbol{b}]} - 1_{[Y_t \leq X_t'\boldsymbol{b}^*]} - [F_t(\boldsymbol{b}) - F_t(\boldsymbol{b}^*)]\}$. As the indicator functions

$1_{[Y_t \leq X_t'\boldsymbol{b}]}, 1_{[Y_t \leq X_t'\boldsymbol{b}^*]}$ and the cumulative density functions $F_t(\boldsymbol{b}), F_t(\boldsymbol{b}^*)$ are functions of bounded

variation, the iid assumption (Assumption 1) and the moment conditions (Assumptions $3'''$ and

12) are sufficient to show that $T^{-1/2} \sum_{t=1}^{T} H(Z_t, \boldsymbol{b})$ is stochastically equicontinuous; that is,

$\sup_{\boldsymbol{b} \in B_T} |T^{-1/2} \sum_{t=1}^{T} H(Z_t, \boldsymbol{b})| \leq o_p(1)$ using Theorems II.2 and II.3 in Andrews (1989).

Theorem 3. *Suppose that Assumptions 1, 2, $3'''$, 4-13 hold.*

(i) Suppose that the conditional quantile model is correctly specified. Then

$$\Sigma_0^{-1/2} T^{1/2} m(\hat{\boldsymbol{b}}_T) \xrightarrow{d} N(0, I_{s \times s})$$

where $\Sigma_0 \equiv \boldsymbol{q}(1 - \boldsymbol{q})(A_0' Q_0^{-1} Q Q_0^{-1} A_0 - A_0' Q_0^{-1} A - A' Q_0^{-1} A_0 + D)$ *with*

$A_0 \equiv E[f_{e|X}(0 \mid X_t) X_t h(X_t)']$, $A \equiv E(X_t h(X_t)')$ *and* $D \equiv E(h(X_t)h(X_t)')$.

(ii) Suppose that the conditional quantile model is correctly specified and that there is no
   conditional heteroaltitudinality in f. Then

$$\Sigma^{-1/2} T^{1/2} m(\hat{\boldsymbol{b}}_T) \xrightarrow{d} N(0, I_{s \times s})$$

where $\Sigma \equiv \boldsymbol{q}(1 - \boldsymbol{q})(D - A'QA)$.

It is straightforward to derive consistent estimators for A, A_0, and D using the plug-in principle. For example,

$$\hat{A}_T \equiv T^{-1} \sum_{t=1}^{T} X_t h(X_t)'$$

$$\hat{A}_{0T} \equiv (2\hat{c}_T T)^{-1} \sum_{t=1}^{T} 1_{[-\hat{c}_T \le \hat{e}_t \le \hat{c}_T]} X_t h(X_t)'$$

$$\hat{D}_T \equiv T^{-1} \sum_{t=1}^{T} h(X_t) h(X_t)'$$

where $\hat{c}_T$ and $\hat{e}_t$ are the same as in the definition of $\hat{Q}_{0T}$ in (4.1). It can be easily shown using Lemma 5 that $\hat{A}_T \overset{p}{\to} A$, $\hat{A}_{0T} \overset{p}{\to} A_0$ and $\hat{D}_T \overset{p}{\to} D$. Therefore, the null of correct specification of the conditional quantile function can be tested using the fact that under the null

$$Tm(\hat{\boldsymbol{b}}_T)' \hat{\Sigma}^{-1} m(\hat{\boldsymbol{b}}_T) \overset{d}{\to} \boldsymbol{c}_s^2$$

where $\hat{\Sigma}$ is a consistent estimator for either Σ_0 or Σ as desired, and which can be constructed using $\hat{Q}_{0T}$, $\hat{Q}_T$, $\hat{A}_{0T}$, $\hat{A}_T$ and $\hat{D}_T$.

6. Monte Carlo Simulations

We conduct simulation experiments to investigate the finite sample properties of our new covariance matrix estimator and to compare it with conventional covariance matrix estimators. For later reference we label the various estimators as follows:

Q-SE1: $\boldsymbol{q}(1-\boldsymbol{q})\hat{f}(0)^{-2}\hat{Q}^{-1}$,

Q-SE2: $\boldsymbol{q}(1-\boldsymbol{q})\hat{Q}_0^{-1}\hat{Q}\hat{Q}_0^{-1}$,

Q-SE3: $\hat{Q}_0^{-1}\hat{V}\hat{Q}_0^{-1}$.

We compare the performance of these alternative estimators in three different set-ups:

[Case 1] the linear conditional quantile model is correctly specified and there is no conditional heteroaltitudinality in the density f,

[Case 2] the linear conditional quantile model is correctly specified but there is conditional heteroaltitudinality in the density f,

[Case 3] the linear conditional quantile model is misspecified and there is conditional heteroaltitudinality in the density f.

We expect that Q-SE1 and Q-SE2 will achieve the best performance in Case 1 and Case 2 respectively, but Q-SE3 will be the winner in the most general case [Case 3]. Bootstrapping the covariance matrix for quantile regressions has also gradually gained popularity. Hence, we also include the bootstrap covariance estimator in our simulation study. We use the design matrix bootstrap covariance estimator used in Buchinsky (1995), defined as

$$\frac{1}{B}\sum_{i=1}^{B}(\tilde{\boldsymbol{b}}_i(\boldsymbol{q})-\hat{\boldsymbol{b}}(\boldsymbol{q}))(\tilde{\boldsymbol{b}}_i(\boldsymbol{q})-\hat{\boldsymbol{b}}(\boldsymbol{q}))'$$

where $\hat{\boldsymbol{b}}(\boldsymbol{q})$ is the $\boldsymbol{q}$-quantile estimator, $\tilde{\boldsymbol{b}}_i(\boldsymbol{q})$ is the ith bootstrap estimator obtained by resampling the pair (y_t, X'_t) and B is the number of bootstrap samples. We set B to 500.

When estimating Q-SE2 and Q-SE3, we need to make a choice for the bandwidth parameter c_T. In our simulations we use three different methods to choose the bandwidth parameter, all of which are derived and discussed in Silverman (1986): two parametric choices based on the standard deviation ($\boldsymbol{s}_e^2$) and the interquartile range (R_e) of the underlying density f_e (the expressions are given by $1.06\boldsymbol{s}_e T^{-1/5}$ and $0.79 R_e T^{-1/5}$ respectively) and one non-parametric choice based on least-squares cross-validation.

Let $\Phi(\cdot)$ be the standard normal cumulative density function. We specify the data generating processes for our cases as follows:

[Case 1] $\quad y_t = 1 + X_{t1} + X_{t2} + \boldsymbol{e}_t$

$\quad\quad\quad\quad \boldsymbol{e}_t \sim iid(-\Phi^{-1}(\boldsymbol{q}), 1)$

[Case 2] $\quad y_t = 1 + X_{t1} + X_{t2} + \boldsymbol{e}_t$

$\quad\quad\quad\quad \boldsymbol{e}_t = \boldsymbol{s}(X_t)\boldsymbol{h}_t$

$\quad\quad\quad\quad \boldsymbol{s}(X_t) = 1 + X_{t1}^2 + X_{t2}^2$

$\quad\quad\quad\quad \boldsymbol{h}_t \sim iid(-\Phi^{-1}(\boldsymbol{q}), 1)$

[Case 3] $\quad y_t = 1 + X_{t1} + X_{t2} + X_{t3} + \sum_{i=1}^{3}\sum_{j=1}^{3} X_{ti}X_{tj} + \boldsymbol{e}_t$

$\quad\quad\quad\quad$ with $\boldsymbol{e}_t$ specified as in Case 2.

In all cases the X_{ti}'s are drawn from the standard normal distribution with cross correlation 0.5; the value for the quantile (q) is set to 0.7. The number of observations and the number of replications are 50 and 1,000 respectively.

For each replication we fit following quantile regression:

$$y_t = \hat{\boldsymbol{b}}_0(\boldsymbol{q}) + \hat{\boldsymbol{b}}_1(\boldsymbol{q})X_{t1} + \hat{\boldsymbol{b}}_2(\boldsymbol{q})X_{t2} + e_t(\boldsymbol{q}).$$

The standard error for each coefficient is calculated using the various methods explained above. We report simulation means of the coefficient estimates and of the standard errors in Table 1. The results for Q-SE2 and Q-SE3 in Table 1 are based on the bandwidth choice using the sample standard deviation, that is $\hat{c}_T = 1.06\hat{\boldsymbol{s}}_e T^{-1/5}$ with $\hat{\boldsymbol{s}}_e^2 = T^{-1}\sum_{t=1}^{T} e_t(\boldsymbol{q})^2$. We obtained quantitatively similar results using other bandwidth choices, so these are omitted.

When the quantile regression model is correctly specified and there is no conditional heteroaltitudinality in the density f [Case 1], all methods to compute standard errors behave similarly except for the bootstrap method, which yields standard errors slightly larger than the other methods. In the last column under the heading "True Std. Dev." we report the simulated standard deviations of the quantile estimates, which closely approximate the true standard deviation of the sampling distribution for the quantile estimator. Not surprisingly, the correct covariance estimator Q-SE1 in Case 1 is smallest and closest to the true standard deviation. It is interesting to note that the efficiency loss caused by using Q-SE2 and Q-SE3 is practically negligible.

In Case 2 where Q-SE2 is the valid covariance estimator, both Q-SE2 and Q-SE3 are reasonably close to the truth. The bootstrap standard errors are giving noticeable over-estimates and Q-SE1 is giving noticeable under-estimates. Finally, we turn to Case 3, where Q-SE3 is the only valid covariance estimator. It is clear from the table that only Q-SE3 is close to the truth. Again, the bootstrap standard errors provide over-estimates. While the other two covariance estimators (Q-SE1 and Q-SE2) both provide under-estimates, Q-SE1 is much worse than Q-SE2. The implication is that, at least in the case considered here, any null hypothesis is more likely to be rejected than it should be when the test is based on Q-SE1 or Q-SE2.

Our simulation study illustrates that: (i) when the linear conditional quantile model is misspecified and there is conditional heteroaltitudinality in the density f, our new covariance

estimator can allow researchers to conduct valid hypothesis tests; (ii) even when these conditions are not satisfied, using our covariance estimator is not likely to cause a serious efficiency loss; and (iii) the bootstrap method to calculate standard errors for quantile regressions should be used with care.

7. Application to Sharpe Style Analysis

In this section, we apply our results to Sharpe style analysis. Following Sharpe's (1988, 1992) seminal work, the Sharpe style regression has become a popular tool to analyze the style of an investment fund. The Sharpe style regression is carried out by regressing fund returns (over a period of time from a fund manager) on various factors mimicking relevant indices. By analyzing the coefficients of the factors, one can understand the style of a fund manager (e.g. style composition, style sensitivity, or style change over time).

As Bassett and Chen (2001) point out, all previous papers have used the method of least squares and hence have concentrated on the relationship between the fund returns and the style factors focused on the mean of the return distribution conditional on the factors. Manager style, however, can be different in different parts of the conditional distribution. For example, a fund manager can change her style when the fund's performance is very good or bad, which corresponds to high quantiles or low quantiles of the conditional distribution. Bassett and Chen (2001) have proposed using the quantile regression method to analyze the style of a fund manager over the entire conditional distribution. They used the returns on the Fidelity Magellan Fund (R_t) over a 5 year sample period (January 1992 – December 1997) with 60 monthly observations. The Russell indices are the factors, which can be classified as follows:

	Large (L)	Small (S)
Growth (G)	Russell 1000 Growth (X_t^{LG})	Russell 2000 Growth (X_t^{SG})
Value (V)	Russell 1000 Value (X_t^{LV})	Russell 2000 Value (X_t^{SV})

The Sharpe style quantile regression equation is then given by

$$R_t = \boldsymbol{a}(\boldsymbol{q}) + \boldsymbol{b}_{LG}(\boldsymbol{q})X_t^{LG} + \boldsymbol{b}_{LV}(\boldsymbol{q})X_t^{LV} + \boldsymbol{b}_{SG}(\boldsymbol{q})X_t^{SG} + \boldsymbol{b}_{SV}(\boldsymbol{q})X_t^{SV} + \boldsymbol{e}_t. \qquad (7.1)$$

On the grounds that for the equity-only funds (e.g. the Magellan Fund or S&P 500 Index), unconstrained and constrained (non-negativity and summing-to-one) cases are usually similar, Bassett and Chen estimate the equation in (7.1) without the non-negativity and summing-to-one constraints. Their findings can be summarized as follows: (i) for the conditional mean, the Magellan fund has an important Large-Value tilt (coefficient 0.69) and otherwise is equally divided between Large-Growth (0.14) and Small-Growth (0.20); (ii) the tilt to Large-Value appears at the other quantiles with an exception being the Quantile $\boldsymbol{q} = 0.1$, where the coefficient for Large-Growth is the largest; and (iii) in all the quantile regressions, most coefficients are not significant due to large standard errors.

We use the same data set with a longer sample period (January 1979 - December 1997), yielding 228 monthly observations. Figure 1 shows a time-series plot of the Magellan Fund. Our belief is that the lack of significance encountered by Bassett and Chen is due to the relatively small number of observations. Since we are not sure about the correctness of the linear conditional quantile specification in (7.1), but would like to keep the linear specification, we calculate standard errors using the various methods explained in Section 6.

The results for the least squares and quantile regressions are reported in Table 2. We take representative values of 0.1, 0.3, 0.5, 0.7 and 0.9 for $\boldsymbol{q}$ in our quantile regressions. For the conditional mean of the distribution (that is, from the least squares regression), the Magellan fund appears to be heavily oriented toward Large-Growth (0.40) and also has an important Large-Value tilt (0.30). The remaining share is equally divided between Small-Growth (0.18) and Small-Value (0.21). In contrast to the findings of Bassett and Chen (2001), the Large-Growth component clearly stands out. This is, however, consistent with their finding that Large-Value orientation is an important component of the style of the Magellan Fund. Further, it is obvious from Figure 2 that the stock market crash in 1987 generated a huge outlier in the returns series. Considering that the least squares estimator is sensitive to outliers, one might like to see how robust the results are given this circumstance. The least absolute deviations (LAD) estimator is a potentially less sensitive alternative (but see Sakata and White (1995, 1998)). The LAD results are reported in the middle of Table 2 ($\boldsymbol{q} = 0.5$). The coefficient for Large-Growth and Small-Growth are almost unchanged, but the coefficient for Large-Value changes from 0.30 to 0.38

while the coefficient for Small-Value has been reduced by half. As we change the value of q, the style for Large-Growth ($\hat{b}_{LG}(q)$) is also gradually decreasing with q while the style for Large-Value ($\hat{b}_{LV}(q)$) becomes more important as q increases except at $q = 0.9$ where there is a sudden drop. The style pattern for Small-Growth ($\hat{b}_{SG}(q)$) is also noticeably changing with q. The tilt to Small-Growth is substantially increasing with q, indicating that the fund tends to invest heavily in Small-Growth stocks when the fund's performance is good, but reduces its share to a statistically insignificant point (when $q = 0.1$, $\hat{b}_{SG}(q)$ is not significant) when the fund's performance is poor. The allocation to Small-Value ($\hat{b}_{SV}(q)$) is decreasing with q except at $q = 0.9$ where it is sharply increasing. In order to see the change in the style against q in detail, we examine a grid of values for q (from 0.1 to 0.9 with 0.1 increment) and plot each quantile estimate $\hat{b}(q)$ with its 95% confidence interval (constructed using Q-SE3) against q. This plot is displayed in Figure 2. The figure confirms our earlier observations.

The pattern of the quantile style (as a function of q) we have found is qualitatively similar to the findings in Bassett and Chen (2001) except for certain large values of q, but we can now provide confidence intervals around the quantile style weights that are robust to the potential misspecification of the conditional quantile function.

In order to examine the potential for misspecification, we apply our quantile specification test using the selection function $h(X_t) = vech(JX_tX_t'J')$ with $J = [0_{4\times1} \ I_{4\times4}]$. The test statistics for the selected values for q and for the three alternatives to compute the bandwidth are given in Table 3. The results are fairly robust to the choice of the bandwidth. The overall conclusion is that for most quantiles we do not have evidence strong enough to reject at the 5% level the null that the linear quantile model in (7.1) is correctly specified. The last row in the table, however, indicates that when $q = 0.9$, the linear specification in (7.1) may be misspecified. It is worth noting again that the standard errors in Table 2 and the confidence intervals in Figure 2 are still valid under this potentially misspecified circumstance.

18

8. Conclusion

We have obtained the asymptotic normality of the quantile estimator for a possibly misspecified model and provided a consistent estimator of the asymptotic covariance matrix. This covariance estimator is misspecification-consistent, that is, it is still valid under misspecification. If researchers confine themselves to a parametric world, then our results are useful when there is uncertainty about correct specification or when one wishes to maintain a model that does not pass a specification test. Although we have restricted our discussion to the linear conditional quantile model for iid data, our methods extend straightforwardly to non-linear conditional quantile models with dependent and possibly heterogeneous data. Investigation of these cases is a promising direction for future research. White (1994, esp. pp. 74-75) provides some consistency results for such cases. See Komunjer (2001) for some asymptotic distribution results in this direction.

Table 1. Simulation Means of 0.7-Quantile Estimates and Standard Errors.

		Quantile Estimates	Q-SE1	Q-SE2	Q-SE3	Bootstrap Std. Errors	True Std. Dev.
Case 1	X_{t1}	1	0.19	0.21	0.21	0.23	0.20
	X_{t2}	1	0.19	0.21	0.20	0.23	0.19
	Constant	0.99	0.19	0.20	0.20	0.22	0.19
Case 2	X_{t1}	1	0.30	0.42	0.41	0.44	0.40
	X_{t2}	0.99	0.30	0.41	0.40	0.43	0.37
	Constant	0.99	0.29	0.32	0.31	0.34	0.28
Case 3	X_{t1}	1.36	1.04	1.24	1.37	1.59	1.35
	X_{t2}	1.32	1.04	1.23	1.36	1.60	1.37
	Constant	5.25	0.89	1.16	1.18	1.40	1.19

Table 2. Mean Style and Quantile Style for Fidelity Magellan Fund

Estimation	Std. Err.	$\hat{b}_{LG}(q)$	$\hat{b}_{LV}(q)$	$\hat{b}_{SG}(q)$	$\hat{b}_{SV}(q)$	$\hat{a}(q)$
Least		**0.40**	**0.30**	**0.18**	**0.21**	**0.45**
Squares	LS-SE1	0.07	0.08	0.06	0.08	0.11
	LS-SE2	0.07	0.08	0.06	0.07	0.11
		0.49	**0.21**	**0.01**	**0.36**	**-1.43**
Quantile	Q-SE1	0.11	0.13	0.09	0.12	0.17
$q = 0.1$	Q-SE2	0.15	0.19	0.10	0.16	0.19
	Q-SE3	0.15	0.21	0.10	0.16	0.19
		0.45	**0.30**	**0.11**	**0.19**	**-0.25**
Quantile	Q-SE1	0.08	0.09	0.07	0.09	0.13
$q = 0.3$	Q-SE2	0.10	0.12	0.08	0.08	0.16
	Q-SE3	0.10	0.12	0.09	0.08	0.16
		0.40	**0.38**	**0.18**	**0.12**	**0.44**
Quantile	Q-SE1	0.08	0.09	0.06	0.09	0.12
$q = 0.5$	Q-SE2	0.07	0.09	0.08	0.09	0.13
	Q-SE3	0.07	0.09	0.08	0.09	0.13
		0.35	**0.39**	**0.24**	**0.10**	**1.07**
Quantile	Q-SE1	0.08	0.10	0.07	0.09	0.13
$q = 0.7$	Q-SE2	0.06	0.08	0.06	0.07	0.14
	Q-SE3	0.06	0.08	0.06	0.07	0.13
		0.27	**0.20**	**0.32**	**0.26**	**2.49**
Quantile	Q-SE1	0.12	0.14	0.10	0.13	0.19
$q = 0.9$	Q-SE2	0.15	0.20	0.09	0.12	0.25
	Q-SE3	0.13	0.20	0.08	0.11	0.25

Note: LS-SE1 = Conventional standard errors for the least squares estimates.

   LS-SE1 = White's heteroskedasticity-consistent standard errors for the least squares estimates.

Table 3. Information Matrix Specification Test Statistics

Quantile (q)	Bandwidth Choice: $1.06s_eT^{-1/5}$	Bandwidth Choice: $0.79R_eT^{-1/5}$	Bandwidth Choice: Least-squares cross-validation
0.1	12.49	15.79	12.40
0.3	14.52	14.42	14.48
0.5	15.62	15.42	15.42
0.7	16.33	16.58	16.58
0.9	22.10	22.13	22.13

Note: The critical value at the 5%-level for the c^2 distribution with 10 degrees of freedom is 18.31.

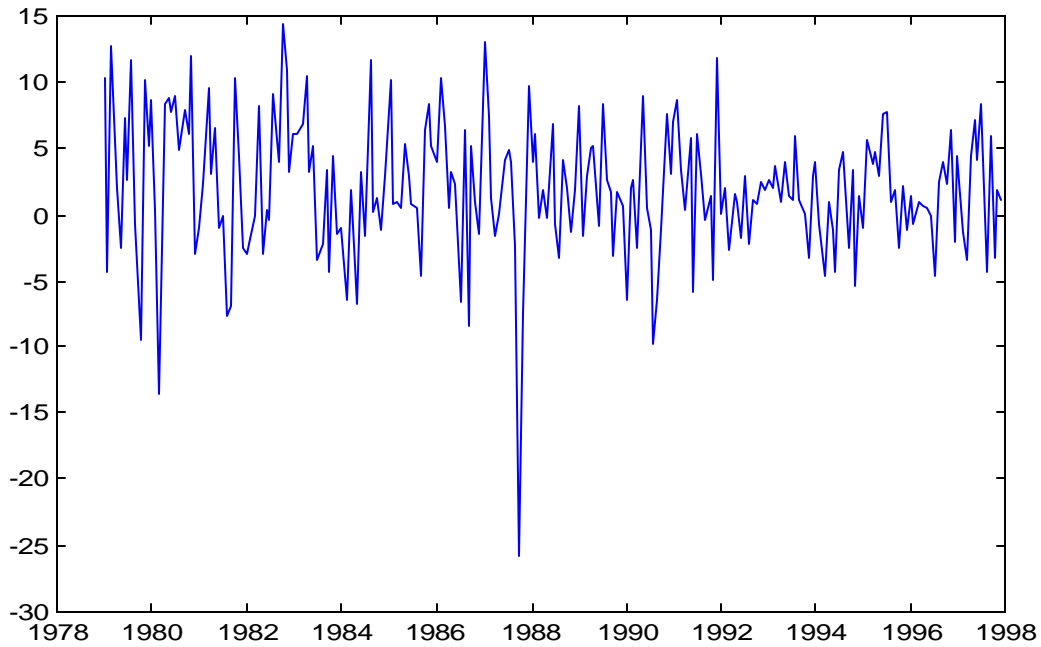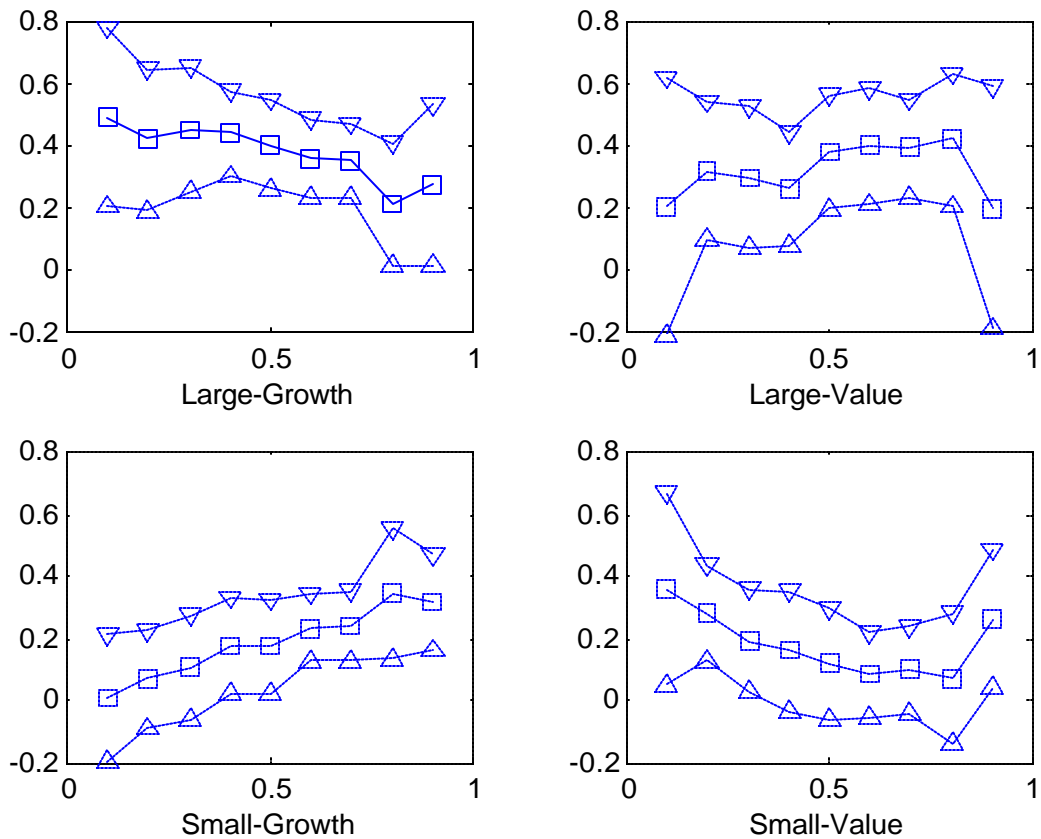Figure 1. Time-series Plot of Fidelity Magellan Fund Monthly Returns

Figure 2. Quantile Style for Fidelity Magellan Fund

Mathematical Appendix

Proof of Lemma 1. We define $Q_T(\boldsymbol{b}) \equiv S_T(\boldsymbol{b}) - S_T(\boldsymbol{b}^*)$ and $\overline{Q}(\boldsymbol{b}) \equiv E(Q_T(\boldsymbol{b}))$. Note that $\overline{Q}(\boldsymbol{b})$ does not depend on T as a result of the iid assumption. Since $S_T(\boldsymbol{b}^*)$ does not depend on $\boldsymbol{b}$,

$$\boldsymbol{b}, \quad \hat{\boldsymbol{b}}_T \in \operatorname*{argmin}_{\boldsymbol{b} \in B} Q_T(\boldsymbol{b}) = T^{-1} \sum_{t=1}^{T} q(Z_t, \boldsymbol{b}) \quad \text{where} \quad q(Z_t, \boldsymbol{b}) \equiv r_q(Y_t - X_t \boldsymbol{b}) - r_q(Y_t - X_t \boldsymbol{b}^*)$$

and $Z_t \equiv (Y_t, X_t')$. In order to apply Lemma 2.2 and Lemma 2.3 in White (1980a), we need to establish the following:

(1) $q(\cdot, \boldsymbol{b})$ is measurable for each $\boldsymbol{b} \in B$.

(2) $q(Z_t, \cdot)$ is continuous on B almost surely.

(3) There exists a measurable function $m : R^{k+1} \to R$ such that

    (i) $|q(z, \boldsymbol{b})| \le m(z)$ for all $\boldsymbol{b} \in B$.

    (ii)   There exists $\boldsymbol{d} > 0$ such that $E|m(Z_t)|^{1+d} \le M < \infty$.

(4) $\overline{Q}(\boldsymbol{b})$ has a unique minimum at $\boldsymbol{b}^*$.

Conditions (1) and (2) are trivially satisfied by inspection of the functional form of q. Let $m(Z_t) \equiv \|X_t\| (\|\overline{\boldsymbol{b}}\| + \|\boldsymbol{b}^*\|)(q+1)$ where $\overline{\boldsymbol{b}}$ is a solution to $\max_{\boldsymbol{b} \in B} \|\boldsymbol{b}\|$. The existence of such a solution is guaranteed by Assumption 5. It is obvious that m is measurable. It is easily seen that $|q(z, \boldsymbol{b})| \le m(z)$ for all $\boldsymbol{b} \in B$. Condition (3)(ii) is satisfied by Assumption $3'$. The last step is to verify condition (4). Let $\boldsymbol{d} \equiv \boldsymbol{b} - \boldsymbol{b}^*$. Since $\overline{Q}(\boldsymbol{b}^*) = 0$ by Assumption 4 and $\overline{Q}(\boldsymbol{b}) = E(q(Z_t, \boldsymbol{b}))$ by Assumption 1, it is sufficient to show that $E(q(Z_t, \boldsymbol{b})) > 0$ for any $\boldsymbol{d} \ne 0$. Note that Assumption 2 and 6 imply that there exists a positive number f_0 such that $|l| < f_0 \Rightarrow f_{e|X}(l \mid x) > f_0$ for all x. Using this property, one can show that

$$E(q(Z_t, \boldsymbol{b})) = E\left[\int_0^{X_t'\boldsymbol{d}} (X_t'\boldsymbol{d} - l) f_{e|X}(l \mid X_t) dl\right]$$

$$\ge f_0 A(\boldsymbol{d})$$

where $A(\boldsymbol{d}) \equiv E\left[\int_0^{X_t'\boldsymbol{d}} (\boldsymbol{1} - X_t'\boldsymbol{d})1_{[-f_0 < \boldsymbol{1} < f_0]}d\boldsymbol{1}\right]$. It is easily checked that $A(\boldsymbol{d}) > 0$ for any $\boldsymbol{d} \neq 0$ in all possible cases: (1) $X_t'\boldsymbol{d} < -f_0$ (2) $-f_0 < X_t'\boldsymbol{d} \leq f_0$ (3) $f_0 < X_t'\boldsymbol{d}$. Therefore, we have the desired result: $\hat{\boldsymbol{b}}_T - \boldsymbol{b}^* = o_p(1)$. ∎

Proof of Lemma 2. It is sufficient to show that $T^{-1/2}\sum_{t=1}^T X_{tj}\boldsymbol{y}_q(Y_t - X_t'\hat{\boldsymbol{b}}_T) = o_p(1)$ for

$j = 1, 2, ..., k$. Let $G_j(a) \equiv \sum_{t=1}^T \boldsymbol{r}_q(Y_t - X_t'\hat{\boldsymbol{b}}_T + X_{tj}a)$ and $H_j(a) \equiv \sum_{t=1}^T X_{tj}\boldsymbol{y}_q(Y_t - X_t'\hat{\boldsymbol{b}}_T + X_{tj}a)$.

Then it can be shown by the definition of $\hat{\boldsymbol{b}}_T$ that $G_j(a)$ achieves its minimum at $a = 0$ and $H_j(a)$ is the left partial derivative of $G_j(a)$. Using these properties, one can show (See Ruppert and Carroll (1980) or Powell (1984) for details) that

$$|H_j(0)| \leq \sum_{t=1}^T |X_{tj}|1_{[Y_t - X_t'\hat{\boldsymbol{b}}_T = 0]}$$

which implies that

$$|T^{-1/2}\sum_{t=1}^T X_{tj}\boldsymbol{y}_q(Y_t - X_t'\hat{\boldsymbol{b}}_T)| \leq T^{-1/2}\sum_{t=1}^T |X_{tj}|1_{[Y_t - X_t'\hat{\boldsymbol{b}}_T = 0]}$$

$$\leq T^{-1/2}\max_{1 \leq t \leq T}\|X_t\| \sum_{t=1}^T 1_{[Y_t - X_t'\hat{\boldsymbol{b}}_T = 0]} .$$

Note that (i) $T^{-1/2}\max_{1 \leq t \leq T}\|X_t\| = o_p(1)$ by Assumption $3'$ and (ii) $\sum_{t=1}^T 1_{[y_t - X_t'\hat{\boldsymbol{b}}_T = 0]} = O_{a.s.}(1)$ by Assumption 2 (See Koenker and Bassett (1978) for details). Hence, the proof is complete. ∎

Proof of Lemma 3 Consider the absolute value of the $(i, j)^{th}$ component of the difference between $Q^* - Q_0$. Using Assumptions $3'''$ and 7, we have

$$|E[(f_{e|X}(\boldsymbol{1}_t^* | X_t) - f_{e|X}(0 | X_t))X_{ti}X_{tj}']| \leq L_0 E[\|X_t\|^3]\|\boldsymbol{b} - \boldsymbol{b}^*\|,$$

which implies that $Q^* = Q_0 + O(\|\boldsymbol{b} - \boldsymbol{b}^*\|)$. Therefore,

$$\boldsymbol{l}(\boldsymbol{b}) - \boldsymbol{l}(\boldsymbol{b}^*) = -Q_0(\boldsymbol{b} - \boldsymbol{b}^*) + O(\|\boldsymbol{b} - \boldsymbol{b}^*\|^2)$$

$$= -Q_0(\boldsymbol{b} - \boldsymbol{b}^*) + o(\|\boldsymbol{b} - \boldsymbol{b}^*\|)$$

which delivers the desired result. ■

Proof of Lemma 4. Define $u(Z_t, \boldsymbol{b}, d) \equiv \sup\limits_{\|\boldsymbol{t}-\boldsymbol{b}\|\leq d} \| X_t[\mathbf{y_q}(Y_t - X_t'\boldsymbol{t}) - \mathbf{y_q}(Y_t - X_t'\boldsymbol{b})] \|$. In order to

invoke Theorem 3 in Huber (1967), we need to verify the following conditions:

(1) $\boldsymbol{b}^*$ is an interior point of the parameter space B.

(2) The first order condition (3.1) is satisfied.

(3) $\hat{\boldsymbol{b}}_T - \boldsymbol{b}^* = o_p(1)$.

(4) For each $\boldsymbol{b} \in B$, $X_t \mathbf{y_q}(Y_t - X_t'\boldsymbol{b})$ is measurable and separable in the sense of Doob (1953).

(5) There exists $\boldsymbol{b}_0 \in B$ such that $\boldsymbol{l}(\boldsymbol{b}_0) = 0$.

(6) There exist $a > 0$ and $d_0 > 0$ such that $\| \boldsymbol{l}(\boldsymbol{b}) \| \geq a \| \boldsymbol{b} - \boldsymbol{b}^* \|$ for $\| \boldsymbol{b} - \boldsymbol{b}^* \| \leq d_0$.

(7) There exist $b > 0$ and $d_0 > 0$ such that $E[u(Z_t, \boldsymbol{b}, d)] \leq bd$ for $\| \boldsymbol{b} - \boldsymbol{b}^* \| + d \leq d_0$.

(8) There exist $c > 0$ and $d_0 > 0$ such that $E[u(Z_t, \boldsymbol{b}, d)^2] \leq cd$ for $\| \boldsymbol{b} - \boldsymbol{b}^* \| + d \leq d_0$.

(9) $E \| X_t \mathbf{y_q}(\boldsymbol{e}_t) \|^2 < \infty$.

First, note that condition (1) is just Assumption 8 and conditions (2) and (3) have been proved in Lemma 1 and Lemma 2 using Assumptions 1, 2, 3′, 4-6. Condition (4) is easily checked using the equivalent definition of separability in Billingsley (1986). This condition ensures that the function $u(\cdot, \boldsymbol{b}, d)$ is measurable. Condition (5) is satisfied given Assumption 4 by letting $\boldsymbol{b}_0 = \boldsymbol{b}^*$. We now verify condition (6). Let a be the smallest eigenvalue of $E[f_{\boldsymbol{e}|X}(\boldsymbol{l}_t^* \,|\, X_t) X_t X_t']$. We can make $E[f_{\boldsymbol{e}|X}(\boldsymbol{l}_t^* \,|\, X_t) X_t X_t']$ sufficiently close to $E[f_{\boldsymbol{e}|X}(0 \,|\, X_t) X_t X_t']$ which is positive-definite by Assumption 9 by choosing a sufficiently small $d_0 > 0$. We choose such $d_0 > 0$ so that a is positive. We have by (3.2)

$$\| \boldsymbol{l}(\boldsymbol{b}) \| = \| Q^*(\boldsymbol{b} - \boldsymbol{b}^*) \| \geq a \| \boldsymbol{b} - \boldsymbol{b}^* \| \text{ for } \| \boldsymbol{b} - \boldsymbol{b}^* \| \leq d_0$$

which verifies condition (6). Some simple algebra gives the inequality

$$u(Z_t, \boldsymbol{b}, d) \leq \| X_t \| 1_{[|e_t - X_t'(\boldsymbol{b} - \boldsymbol{b}^*)| \ \leq \ \|X_t\|d]}$$

which implies by the law of iterated expectations and Assumption 10 that

$$E[u(Z_t, \boldsymbol{b}, d)] \leq 2 f_1 E[\| X_t \|^2] d.$$

Let $b \equiv 2f_1 E[\| X_t \|^2]$. Then b is positive and finite by Assumption $3'$ which verifies condition (7). Condition (8) can be verified in a similar fashion by letting $c \equiv 2f_1 E[\| X_t \|^3] > 0$, which is finite by Assumption $3'''$. Note that $E \| X_t \boldsymbol{y_q}(\boldsymbol{e}_t) \|^2 = E[\boldsymbol{y_q}(\boldsymbol{e}_t)^2 X_t' X_t] \le E[X_t' X_t] < 0$ by Assumption $3'$. All the conditions in Theorem 3 in Huber (1967) are thus satisfied. We therefore have the desired result:

$$T^{-1/2} \sum_{t=1}^{T} X_t \boldsymbol{y_q}(\boldsymbol{e}_t) + T^{1/2} \boldsymbol{l}(\hat{\boldsymbol{b}}_T) = o_p(1). \quad \blacksquare$$

Proof of Theorem 1. Using the Lindeberg-Levy CLT, Assumptions 1, 3, 4 and 11 imply that

$$T^{-1/2} \sum_{t=1}^{T} X_t \boldsymbol{y_q}(\boldsymbol{e}_t) \overset{d}{\to} N(0,V). \tag{A.1}$$

Lemma 3, Lemma 4 and (A.1) together imply that

$$T^{1/2}(\hat{\boldsymbol{b}}_T - \boldsymbol{b}^*) \overset{d}{\to} N(0,C)$$

where $C \equiv Q_0^{-1} V Q_0^{-1}$. $\quad \blacksquare$

Proof of Lemma 5. Consider $Q_T \equiv (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[-c_T \le e_t \le c_T]} X_t X_t'$. Using the mean value theorem, one can show that $E(Q_T) = E\left[T^{-1} \sum_{t=1}^{T} f_{e|X}(\boldsymbol{x}_T \mid X_t) X_t X_t' \right]$ where $-c_T \le \boldsymbol{x}_T \le c_T$. Hence, $\boldsymbol{x}_T = o(1)$ by Assumption 12(ii). Let $Q_0 \equiv E\left[T^{-1} \sum_{t=1}^{T} f_{e|X}(0 \mid X_t) X_t X_t' \right]$. Then it is easily checked that $| E(Q_T) - Q_0 | \to 0$ by the triangle inequality and Assumptions $3'$ and 7. Using a LLN for double arrays (e.g. Theorem 2 in Andrews (1989)), we have that $Q_T \overset{p}{\to} E(Q_T)$. Therefore, we have that $Q_T \overset{p}{\to} Q_0$. The rest of the proof is carried out in two steps:

(1) $| \tilde{Q}_{0T} - Q_T | \overset{p}{\to} 0$ where $\tilde{Q}_{0T} \equiv (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[-\hat{c}_T \le \hat{e}_t \le \hat{c}_T]} X_t X_t'$

(2) $| \hat{Q}_{0T} - \tilde{Q}_{0T} | \overset{p}{\to} 0$.

The $(i, j)^{th}$ element of $| \tilde{Q}_{0T} - Q_T |$ is given by

$$| (2c_T T)^{-1} \sum_{t=1}^{T} (1_{[-\hat{c}_T \le \hat{e}_t \le \hat{c}_T]} - 1_{[-c_T \le e_t \le c_T]}) X_{ti} X_{tj} |$$

$$\le (2c_T T)^{-1} \sum_{t=1}^{T} (1_{[|e_t + c_T| \le d_T]} + 1_{[|e_t - c_T| \le d_T]}) | X_{ti} || X_{tj} | \equiv U_{1T} + U_{2T}$$

where $d_T \equiv \| X_t \| \| \hat{\boldsymbol{b}}_T - \boldsymbol{b}^* \| + | \hat{c}_T - c_T |$ (by the fact that $| 1_{[x \le 0]} - 1_{[y \le 0]} | \le 1_{[|x| \le |x-y|]}$),

$U_{1T} \equiv (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_t - c_T| \le d_T]} | X_{ti} || X_{tj} |$, and $U_{2T} \equiv (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_t + c_T| \le d_T]} | X_{ti} || X_{tj} |$.

First, we show that $U_{1T} \overset{p}{\to} 0$. For this, let $\boldsymbol{h} > 0$ and consider

$P(U_{1T} > \boldsymbol{h})$

$$= P((2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_t - c_T| \le d_T]} | X_{ti} || X_{tj} | > \boldsymbol{h})$$

$$\equiv P(A) \quad (\text{where } A \equiv [(2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_t - c_T| \le d_T]} | X_{ti} || X_{tj} | > \boldsymbol{h}])$$

$$\le P(A \cap B \cap C) + P(C^c) + P(B^c) \quad \text{for any events } B \text{ and } C.$$

Now let $B \equiv \{c_T^{-1} \| \hat{\boldsymbol{b}}_T - \boldsymbol{b}^* \| \le z\}$ for a constant $z > 0$ and $C \equiv \{c_T^{-1} | \hat{c}_T - c_T | \le z\}$. Note that

as $T \to \infty$, (1) $P(B^c) \to 0$ by Assumption 12(iii) and $T^{1/2} \| \hat{\boldsymbol{b}}_T - \boldsymbol{b}^* \| = O_p(1)$; and (2)

$P(C^c) \to 0$ by Assumption 5(i). Now

$P(A \cap B \cap C)$

$$\le (2\boldsymbol{h} c_T T)^{-1} \sum_{t=1}^{T} E\left[\int_{-(\|X_t\| + 1) z c_T + c_T}^{(\|X_t\| + 1) z c_T + c_T} f_{e|X} (\boldsymbol{l} | X_t) d\boldsymbol{l} \ | X_{ti} || X_{tj} | \right] \quad \text{(by the Markov inequality)}$$

$$\le (2\boldsymbol{h} c_T T)^{-1} \sum_{t=1}^{T} E\left[\int_{-(\|X_t\| + 1) z c_T + c_T}^{(\|X_t\| + 1) z c_T + c_T} f_1 d\boldsymbol{l} \ | X_{ti} || X_{tj} | \right] \quad \text{(by Assumption 10)}$$

$$= z f_1 \boldsymbol{h}^{-1} E[(\| X_t \| + 1) | X_{ti} || X_{tj} |] < \infty \quad \text{(by Assumptions 1 and 3''').}$$

We can choose z arbitrarily small, so $P(U_{1T} > \boldsymbol{h}) \to 0$ which implies that $U_{1T} \overset{p}{\to} 0$ because

$U_{1T} \ge 0$. It can be shown in the same fashion that $U_{2T} \overset{p}{\to} 0$, which completes the first step. To

show the second step, consider $\hat{Q}_{0T} - \tilde{Q}_{0T} = \left(\dfrac{c_T}{\hat{c}_T} - 1\right)\tilde{Q}_{0T}$. Note that $\tilde{Q}_{0T} = O_p(1)$ and

$\left(\dfrac{c_T}{\hat{c}_T} - 1\right) = o(1)$ by Assumption 5(i). Hence, the second step follows, which delivers the desired

result: $\hat{Q}_{0T} \xrightarrow{p} Q_0$. ∎

Proof of Lemma 6. Since the proof is quite similar to the proof of Lemma 5, we do not provide

the details. Let $V_T \equiv T^{-1}\displaystyle\sum_{t=1}^{T} \mathbf{y_q}(\mathbf{e}_t)^2 X_t X_t'$. By the law of large numbers for iid random variables,

we have $V_T \xrightarrow{p} V$ by Assumption 1 and 3″. Next we show that $\hat{V}_T - V_T \xrightarrow{p} 0$. Consider the

$(i, j)^{th}$ element of $|\hat{V}_T - V_T|$:

$$|T^{-1}\sum_{t=1}^{T}[\mathbf{y_q}(\hat{\mathbf{e}}_t)^2 - \mathbf{y_q}(\mathbf{e}_t)^2]X_{ti}X_{tj}|$$

$$\leq (\mathbf{q}+1)^2 T^{-1}\sum_{t=1}^{T}\mathbf{1}_{[|\mathbf{e}_t| \leq d_T]} |X_{ti}||X_{tj}|,$$

(where $d_T \equiv ||X_t||||\hat{\mathbf{b}}_T - \mathbf{b}^*||$)

by the triangle inequality and the Cauchy-Schwarz inequality.

Let $U_T \equiv T^{-1}\displaystyle\sum_{t=1}^{T}\mathbf{1}_{[|\mathbf{e}_t| \leq d_T]}|X_{ti}||X_{tj}|$. By the same argument as in the proof of Lemma 5, one

can show that $U_T \xrightarrow{p} 0$ using Assumptions 1, 3‴, 10, and $T^{1/2}||\hat{\mathbf{b}}_T - \mathbf{b}^*|| = O_p(1)$. ∎

Proof of Lemma 7. First, note that $\mathbf{s}_u^2 \equiv \mathbf{q}(1-\mathbf{q})$. Hence, $E[(u_t^2 - \mathbf{s}_u^2)X_t X_t']$

$$= E[(u_t^2 - \mathbf{q}(1-\mathbf{q}))X_t X_t']$$

$$= E[cE(u_t \mid X_t)X_t X_t'] \text{ where } c \equiv 2\mathbf{q} - 1,$$

which is zero because of correct specification assumption, $E(u_t \mid X_t) = 0$. ∎

Proof of Theorem 3. Since $T^{1/2}(\hat{\mathbf{b}}_T - \mathbf{b}^*) = O_p(1)$, Lemma 8 implies that

$$T^{1/2}m(\hat{\boldsymbol{b}}_T) = T^{1/2}m(\boldsymbol{b}^*) + T^{-1/2}\sum_{t=1}^{T}h(X_t)[F_t(\hat{\boldsymbol{b}}_T) - F_t(\boldsymbol{b}^*)] + o_p(1). \qquad\qquad (A.2)$$

It is straightforward to show that

$$T^{1/2}m(\boldsymbol{b}^*) = T^{-1/2}\sum_{t=1}^{T}h(X_t)\boldsymbol{y_q}(\boldsymbol{e}_t)$$

$$T^{-1/2}\sum_{t=1}^{T}h(X_t)[F_t(\hat{\boldsymbol{b}}_T) - F_t(\boldsymbol{b}^*)] = T^{-1}\sum_{t=1}^{T}f_{\boldsymbol{e}|X}(0\,|\,X_t)h(X_t)X_t'T^{1/2}(\hat{\boldsymbol{b}}_T - \boldsymbol{b}^*) + o_p(1).$$

$$= A_0'Q_0^{-1}T^{-1/2}\sum_{t=1}^{T}X_t\boldsymbol{y_q}(\boldsymbol{e}_t) + o_p(1).$$

Plugging these expressions into (A.2) and collecting terms gives

$$T^{1/2}m(\hat{\boldsymbol{b}}_T) = T^{-1/2}\sum_{t=1}^{T}a_t\boldsymbol{y_q}(\boldsymbol{e}_t) + o_p(1), \qquad\qquad (A.3)$$

where $a_t \equiv A_0'Q_0^{-1}X_t - h(X_t)$. Under the assumption that the conditional quantile model is correctly specified, the Lindeberg-Levy CLT delivers that

$$T^{-1/2}\sum_{t=1}^{T}a_t\boldsymbol{y_q}(\boldsymbol{e}_t) \xrightarrow{d} N(0,\Sigma_a^2)$$

where $\Sigma_0 \equiv \boldsymbol{q}(1-\boldsymbol{q})(A_0'Q_0^{-1}QQ_0^{-1}A_0 - A_0'Q_0^{-1}A - A'Q_0^{-1}A_0 + D)$, which completes the proof of (i). Furthermore, if there is no conditional heteroaltitudinality in f, then it can be shown that $A_0 = f(0)A$ as well as that $Q_0 = f(0)Q$. Then the asymptotic variance Σ_0 in (i) simplifies to

$$\Sigma \equiv \boldsymbol{q}(1-\boldsymbol{q})(D - A'QA),$$

which completes the proof of (ii). ∎

References

Andrews, D.W.K. (1989), "Asymptotics for Semiparametric Econometric Models: II. Stochastic Equicontinuity and Nonparametric Kernel Estimation," Cowles Foundation Discussion Paper No. 909R.

Amemiya, T. (1982), "Two Stage Least Absolute Deviation Estimators," *Econometrica*, 50, 689-711.

Bai, J. (1995), "Least Absolute Deviation Estimation of A Shift," *Econometric Theory*, 11, 403-436.

Bassett, G. and H. L. Chen (2001), "Portfolio Style: Return-Based Attribution Using Regression Quantiles," *Empirical Economics*, 26, 293-305.

Bassett, G. and R. Koenker (1978), "Asymptotic Theory of Least Absolute Error Regression," *Journal of the American Statistical Association*, 73, 618-622.

Bhattacharya, P. K. and A. K. Gangopadhyay (1990), "Kernel and Nearest-Neighbor Estimation of a Conditional Quantile," *The Annals of Statistics*, 18, 1400-1415.

Billingsley, P. (1986). *Probability and Measure.* New York: John Wiley & Sons.

Bloomfield, P. and W. L. Steiger (1983). *Least Absolute Deviations: Theory, Applications, and Algorithms.* Boston: Birkhauser.

Buchinsky, M. (1995), "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study," *Journal of Econometrics*, 68, 303-338.

Buchinsky, M. (1998), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *The Journal of Human Resources*, 1, 88-126.

Buchinsky, M. and J. Hahn (1998), "An Alternative Estimator for the Censored Quantile Regression Model," *Econometrica*, 66, 653-671.

Doob, J. L. (1953). S*tochastic Processes.* New York: John Wiley & Sons.

Jureckova. J. (1977), "Asymptotic Relations of M-Estimates and R-Estimates in Linear Regression Models," *The Annals of Statistics*, 5, 464-472.

Hahn, J. (1995), "Bootstrapping Quantile Regression Estimators," *Econometric Theory*, 11, 105-121.

Herce, M. A. (1996), "Asymptotic Theory of LAD Estimation in a Unit Root Process with Finite Variance Errors," *Econometric Theory*, 12, 129-153.

Horowitz, J. L. (1998), "Bootstrap Methods for Median Regression Models," *Econometrica*, 66, 1327-1351.

Huber, P. J. (1967), "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 221-233.

Koenker, R. and G. Bassett (1978), "Regression Quantiles," *Econometrica*, 46, 33-50.

Koenker, R. and G. Bassett (1982), "Robust Tests for Heteroskedasticity Based on Regression Quantiles," *Econometrica*, 50, 43-61.

Koenker, R. and Q. Zhao (1996), "Conditional Quantile Estimation and Inference for ARCH Models," *Econometric Theory*, 12, 793-813.

Komunjer, I. (2001), "The 'a-Quantile' Distribution Function and its Applications to Financial Modeling," HEC School of Management Department of Finance Working Paper.

MacKinnon, J. G. and H. White (1985), "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics,* 29, 305-325.

Phillips, P. C. B. (1991), "A Shortcut to LAD Estimator Asymptotics," *Econometric Theory,* 7, 450-464.

Powell, J. (1983), "The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators," *Econometrica*, 51, 1569-1575.

Powell, J. (1984), "Least Absolute Deviations Estimation for the Censored Regression Model," *Journal of Econometrics*, 25, 303-25.

Powell, J. (1986), "Censored Regression Quantiles," *Journal of Econometrics*, 32, 143-155.

Pollard, D. (1985), "New Ways to Prove Central Limit Theorems," *Econometric Theory*, 1, 295-314.

Pollard, D. (1991), "Asymptotics for Least Absolute Deviation Regression Estimators," *Econometric Theory*, 7, 186-199.

Ruppert. D. and R. J. Carroll (1980), "Trimmed Least Squares Estimation in the Linear Model," *Journal of the American Statistical Association*, 75, 828-838.

Sharpe, W.F. (1988), "Determining a Fund's Effective Asset Mix," *Investment Management Review*, December, 59-69.

Sharpe, W.F. (1992), "Asset Allocation: Management Style and Performance Measurement," *Journal of Portfolio Management,* 18, 7-19.

Sheather, S. J. and J. S. Marron (1990), "Kernel Quantile Estimators," *Journal of the American Statistical Association*, 85, 410-416.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

Sakata, S. and H. White (1995), "An Alternative Definition of Finite Sample Breakdown Point with Applications to Regression Model Estimators," *Journal of the American Statistical Association*, 90, 1099-1106.

Sakata, S. and H. White (1998), "High Breakdown Point Conditional Dispersion Estimation with Application to S&P 500 Daily Returns Volatility," *Econometrica,* 66, 529-567.

Weiss, A. A. (1990), "Least Absolute Error Estimation in The Presence of Serial Correlation," *Journal of Econometrics*, 44, 127-158.

White, H. (1980a), "Nonlinear Regression on Cross-Section Data," *Econometrica*, 48, 721-746.

White, H. (1980b), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

White, H. (1992), "Nonparametric Estimation of Conditional Quantiles using Neural Networks," In H. White (ed.), *Artificial Neural Networks*: *Approximation and Learning Theory*, Oxford: Blackwell, pp. 191-205.

White, H. (1994). *Estimation, Inference and Specification Analysis.* New York: Cambridge University Press.

Zheng, J. X. (1998), "A Consistent Nonparametric Test of Parametric Regression Models Under Conditional Quantile Restrictions," *Econometric Theory*, 14, 123-238.