# An Encompassing Test for Non-Nested Quantile Regression Models

**Chung-Ming Kuan**
Department of Finance
National Taiwan University

**Hsin-Yi Lin**
Department of Economics
National Chengchi University

## Abstract

We propose an encompassing test for non-nested linear quantile regression models and show that it has an asymptotic $\chi^2$ distribution. It is also shown that the proposed test is a regression rank score test in a comprehensive model under conditional homogeneity. Our simulation results indicate that the proposed test performs very well in finite samples.

**JEL classification:** C12, C52
**Keywords:** encompassing test; non-nested model; quantile regression; rank score test.

† Author for correspondence: Hsin-Yi Lin, Department of Economics, National Chengchi University, Taipei 116, Taiwan. Email: linh@nccu.edu.tw; tel.: +886 2 29393091 ext. 51547; fax: +886 2 29390344.

†† This paper is part of the project "Advancement of Research on Econometric Methods and Applications" (AREMA). We thank Roger Koenker, Yu-Chin Hsu, and a referee for very useful comments and suggestions. Lin wishes to express her sincere gratitude to the Ministry of Education for supporting her postdoctoral research and the University of Illinois, Urbana-Champaign, for arranging her visit. C.-M. Kuan and H.-Y. Lin gratefully acknowledge research supports from the National Science Council of the Republic of China (NSC 93-2415-H-001-028 for Kuan and NSC 94-2415-H-004-022 for Lin).

# 1 Introduction

There are two leading approaches to constructing non-nested tests: the comprehensive-model approach (Atkinson, 1970) and the encompassing approach (Mizon, 1984; Mizon and Richard, 1986). The former bases tests on an artificial nesting model that includes competing models as special cases, e.g., Davidson and MacKinnon (1981), Fisher and McAleer (1981), MacKinnon *et al.* (1983), and Santos Silva (2001). The latter compares a statistic of the alternative model and its pseudo-true value, e.g., Gouriéroux *et al.* (1983), Wooldrdige (1990), Smith (1992), Chen and Kuan (2002, 2007), and Ramalho and Simth (2002). Yet, Bontemps and Mizon (2008) showed that there is implicitly a comprehensive model underlying each test obtained from the encompassing approach.

Although the encompassing approach is valid in general, the existing encompassing tests are not readily applied to the models for which the objective function or the estimating function is not smooth, such as the quantile regression (QR) of Koenker and Bassett (1978). In this paper we follow the encompassing approach to obtain the conditional quantile encompassing (CQE) test for non-nested, linear QR models. We show that the CQE test has an asymptotic $\chi^2$ distribution and is asymptotically equivalent to the regression rank score test of Gutenbrunner *et al.* (1993) in a comprehensive model under conditional homogeneity. Our simulation results indicate that the proposed test performs very well in finite samples.

The paper proceeds as follows. We derive the CQE test and its limiting distribution in Section 2. We show the CQE test is a rank score test in Section 3. Simulation results are reported in Section 4. Section 5 concludes this paper.

# 2 The Conditional Quantile Encompassing Test

Let $(y_i, \boldsymbol{x}_i', \boldsymbol{z}_i')'$, $i = 1, \ldots, n$, be independent random vectors, with $\boldsymbol{x}_i$ a $p \times 1$ vector and $\boldsymbol{z}_i$ a $q \times 1$ vector. Let $Q_{y_i|\mathcal{F}_i}(\tau)$, $\tau \in (0, 1)$, denote the $\tau$-th conditional quantile function of $y_i$ given $\mathcal{F}_i$, the information set generated by $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$. We want to test the QR model:

$$M_0 \colon Q_{y_i|\mathcal{F}_i}(\tau) = \boldsymbol{x}_i'\boldsymbol{\beta}(\tau), \quad i = 1, \ldots, n,$$

against the alternative specification:

$$M_1 \colon Q_{y_i|\mathcal{F}_i}(\tau) = \boldsymbol{z}_i'\boldsymbol{\gamma}(\tau), \quad i = 1, \ldots, n,$$

where $\boldsymbol{\beta}(\tau)$ and $\boldsymbol{\gamma}(\tau)$ are unknown parameter vectors. Let $\rho_\tau(u) = u[\tau - \mathbf{1}(u < 0)]$, with $\mathbf{1}(A)$ the indicator function of the event $A$. The QR estimator of $\boldsymbol{\beta}(\tau)$ in model $M_0$, denoted as $\hat{\boldsymbol{\beta}}_n(\tau)$, minimizes $n^{-1}\sum_{i=1}^n \rho_\tau(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})$. The QR estimator of $\boldsymbol{\gamma}(\tau)$ in model $M_1$, denoted as $\hat{\boldsymbol{\gamma}}_n(\tau)$, minimizes $n^{-1}\sum_{i=1}^n \rho_\tau(y_i - \boldsymbol{z}_i'\boldsymbol{\gamma})$. The parameter estimates may be computed via a linear programming algorithm (Koenker and Bassett, 1978).

The existing encompassing tests can not be directly applied to test non-nested QR models. A parameter encompassing test requires evaluating the pseudo-true value (i.e., the probability limit under the null hypothesis) of the parameter estimator for the alternative model. This is, however, a formidable task because the QR estimator does not have an analytic form. Chen and Kuan (2002) propose the pseudo-true score encompassing (PSE) test that compares the score function of the alternative model and its pseudo-true value. The PSE test is appealing because the functional form of score function is usually available and its pseudo-true value is relative easy to obtain. Yet, the PSE test is derived when the objective (likelihood) function and score function are "smooth." This approach can be extended to construct an encompassing test for QR models, even when $\rho_\tau(u)$ in the QR objective function is not differentiable at $u = 0$.

The first-order derivative of the QR objective function for model $M_0$, except at $y_i = \boldsymbol{x}_i'\boldsymbol{\beta}$, is $\boldsymbol{s}_\tau^0(\boldsymbol{\beta}) = n^{-1}\sum_{i=1}^n \boldsymbol{x}_i[\tau - \mathbf{1}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta} < 0)]$. This function is understood as an (approximate) estimating function for the QR model $M_0$ because it, when evaluated at $\hat{\boldsymbol{\beta}}_n(\tau)$, yields the "asymptotic first-order condition" (see, e.g., Otsu, 2008): $n^{1/2}\boldsymbol{s}_\tau^0\big(\hat{\boldsymbol{\beta}}_n(\tau)\big) = o_{\mathrm{IP}}(1)$. Similarly,

$$\boldsymbol{s}_\tau^1(\boldsymbol{\gamma}) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{z}_i[\tau - \mathbf{1}(y_i - \boldsymbol{z}_i'\boldsymbol{\gamma} < 0)]$$

is an estimating function for the alternative model $M_1$, and $n^{1/2}\boldsymbol{s}_\tau^0\big(\hat{\boldsymbol{\gamma}}_n(\tau)\big) = o_{\mathrm{IP}}(1)$. The estimating function plays the role of the score function under the likelihood framework.

In the light of Chen and Kuan (2002), it is natural to base an encompassing test on the difference between $\boldsymbol{s}_\tau^1(\boldsymbol{\gamma})$ and its pseudo-true value. Under the null, model $M_0$ is the $\tau$-th conditional quantile function and satisfies the conditional moment restriction:

$$\mathrm{IE}_0[\tau - \mathbf{1}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0)|\mathcal{F}_i] = \tau - F_{y_i|\mathcal{F}_i}(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)) = 0, \quad \text{a.s.,} \tag{1}$$

where $\mathrm{IE}_0$ denotes the expectation under the null, and $F_{y_i|\mathcal{F}_i}$ is the distribution function of $y_i$ conditional on $\mathcal{F}_i$ (with the conditional density $f_{y_i|\mathcal{F}_i}$). By the restriction (1) and the law of iterated expectations,

$$\mathrm{IE}_0[\boldsymbol{s}_\tau^1(\boldsymbol{\gamma})] = \frac{1}{n}\mathrm{IE}_0\sum_{i=1}^n \boldsymbol{z}_i[\mathbf{1}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0) - \mathbf{1}(y_i - \boldsymbol{z}_i'\boldsymbol{\gamma} < 0)].$$

The limit of this function is the pseudo-true value of $\boldsymbol{s}_\tau^1(\boldsymbol{\gamma})$, and its sample counterpart is

$$\hat{\boldsymbol{\zeta}}_n(\boldsymbol{\beta}(\tau), \boldsymbol{\gamma}) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{z}_i\big[\mathbf{1}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0) - \mathbf{1}(y_i - \boldsymbol{z}_i'\boldsymbol{\gamma} < 0)\big].$$

The proposed CQE test is based on:

$$\sqrt{n}\big[\boldsymbol{s}_\tau^1\big(\hat{\boldsymbol{\gamma}}_n(\tau)\big) - \hat{\boldsymbol{\zeta}}_n\big(\hat{\boldsymbol{\beta}}_n(\tau), \hat{\boldsymbol{\gamma}}_n(\tau)\big)\big] = \frac{1}{\sqrt{n}}\sum_{i=1}^n \boldsymbol{z}_i\big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0\big)\big]. \tag{2}$$

The function in (2) will be denoted as $\hat{\boldsymbol{\psi}}_n\big(\hat{\boldsymbol{\beta}}_n(\tau)\big)$ and is analogous to the basic ingredient of the conditional mean encompassing (CME) test of Wooldridge (1990), a special case of the PSE test for non-nested (non)linear models. Note that a test based on (2) is also a test of the conditional moment restriction (1). When $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ have $r$ variables in common ($r < p, q$), let $\boldsymbol{z}_i^\dagger$ denote the sub-vector of the common variables in $\boldsymbol{z}_i$ and $\boldsymbol{z}_i^*$ the sub-vector of remaining $q - r$ variables. We may ignore $\boldsymbol{z}_i^\dagger$ and base the CQE test on

$$\hat{\boldsymbol{\psi}}_n^*\big(\hat{\boldsymbol{\beta}}_n(\tau)\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{z}_i^* \big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0\big)\big], \tag{3}$$

since $n^{-1/2}\sum_{i=1}^n \boldsymbol{z}_i^\dagger\big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0\big)\big] = o_{\mathbb{P}}(1)$ by the asymptotic first-order condition.

As shown in Appendix, (3) can be expressed in terms of the true parameter $\boldsymbol{\beta}(\tau)$:

$$\hat{\boldsymbol{\psi}}_n^*(\hat{\boldsymbol{\beta}}_n(\tau)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(\boldsymbol{z}_i^* - \boldsymbol{M}_{fzx}\boldsymbol{M}_{fxx}^{-1}\boldsymbol{x}_i\big)\big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0\big)\big] + o_{\mathbb{P}}(1), \tag{4}$$

with $\boldsymbol{M}_{fzx} = \operatorname{plim}_n \sum_{i=1}^n f_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big)\boldsymbol{z}_i^*\boldsymbol{x}_i'/n$ and $\boldsymbol{M}_{fxx} = \operatorname{plim}_n \sum_{i=1}^n f_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big)\boldsymbol{x}_i\boldsymbol{x}_i'/n$ (assuming these limits exist). Under the null hypothesis, $\mathbf{1}\big(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0\big)$ has conditional mean $\tau$ and conditional variance $\tau(1-\tau)$. When the data are independent random variables obeying the Lindeberg-Feller central limit theorem (e.g., White, 1999, Theorem 5.6), we have

$$\hat{\boldsymbol{\psi}}_n^*\big(\hat{\boldsymbol{\beta}}_n(\tau)\big) \xrightarrow{D} \mathcal{N}\big(\mathbf{0}, \tau(1-\tau)\boldsymbol{V}_{o,\tau}\big),$$

where $\boldsymbol{V}_{o,\tau} = \lim_{n\to\infty} n^{-1}\mathbb{E}_o \sum_{i=1}^n (\boldsymbol{z}_i^* - \boldsymbol{M}_{fzx}\boldsymbol{M}_{fxx}^{-1}\boldsymbol{x}_i)(\boldsymbol{z}_i^* - \boldsymbol{M}_{fzx}\boldsymbol{M}_{fxx}^{-1}\boldsymbol{x}_i)'$. Letting $\widehat{\boldsymbol{V}}_{n,\tau}$ be a consistent estimator of $\boldsymbol{V}_{o,\tau}$, the proposed CQE test is

$$\boldsymbol{\Psi}_{n,\tau}^* = \frac{1}{\tau(1-\tau)} \hat{\boldsymbol{\psi}}_n^*\big(\hat{\boldsymbol{\beta}}_n(\tau)\big)' \, \widehat{\boldsymbol{V}}_{n,\tau}^{-1} \, \hat{\boldsymbol{\psi}}_n^*\big(\hat{\boldsymbol{\beta}}_n(\tau)\big), \tag{5}$$

which has $\chi^2(q-r)$ distribution under the null.

The matrices $\boldsymbol{M}_{fzx}$ and $\boldsymbol{M}_{fxx}$ involve conditional density and may be estimated by nonparametric estimators, such as the "Powell sandwich" of Powell (1991); see also Koenker (2005) for other estimators. Let $\widehat{\boldsymbol{M}}_{fzx}$ and $\widehat{\boldsymbol{M}}_{fxx}$ denote their consistent estimators. Using the matrix notations $\boldsymbol{X}$ (an $n \times p$ matrix with the $i$-th row $\boldsymbol{x}_i'$) and $\boldsymbol{Z}^*$ (an $n \times (q-r)$ matrix with the $i$-th row $\boldsymbol{z}_i^{*\prime}$), a consistent estimator $\widehat{\boldsymbol{V}}_{n,\tau}$ is:

$$\frac{1}{n}\big[\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^* - (\boldsymbol{Z}^{*\prime}\boldsymbol{X})\widehat{\boldsymbol{M}}_{fxx}^{-1}\widehat{\boldsymbol{M}}_{fzx}' - \widehat{\boldsymbol{M}}_{fzx}\widehat{\boldsymbol{M}}_{fxx}^{-1}(\boldsymbol{X}'\boldsymbol{Z}^*) + \widehat{\boldsymbol{M}}_{fzx}\widehat{\boldsymbol{M}}_{fxx}^{-1}(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{M}}_{fxx}^{-1}\widehat{\boldsymbol{M}}_{fzx}'\big].$$

Under conditional homogeneity: $f_{y_i|\mathcal{F}_i} = f_{y_i}$, the test (5) can be easily implemented. First note that (4) becomes

$$\hat{\boldsymbol{\psi}}_n^*(\hat{\boldsymbol{\beta}}_n(\tau)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\boldsymbol{z}_i^* - \boldsymbol{M}_{zx}\boldsymbol{M}_{xx}^{-1}\boldsymbol{x}_i)\big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0\big)\big] + o_{\mathbb{P}}(1),$$

with $\boldsymbol{M}_{zx} = \text{plim}_n \sum_{i=1}^n \boldsymbol{z}_i^* \boldsymbol{x}_i'/n$ and $\boldsymbol{M}_{xx} = \text{plim}_n \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i'/n$. Then, the asymptotic normality of $\hat{\boldsymbol{\psi}}_n^*(\hat{\boldsymbol{\beta}}_n(\tau))$ holds with

$$\boldsymbol{V}_{o,\tau} = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_o \sum_{i=1}^n (\boldsymbol{z}_i^* - \boldsymbol{M}_{zx}\boldsymbol{M}_{xx}^{-1}\boldsymbol{x}_i)(\boldsymbol{z}_i^* - \boldsymbol{M}_{zx}\boldsymbol{M}_{xx}^{-1}\boldsymbol{x}_i)',$$

which does not involve density function. A consistent estimator of $\boldsymbol{V}_{o,\tau}$ is easily computed as $\widehat{\boldsymbol{V}}_{n,\tau} = \boldsymbol{Z}^{*\prime}[\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{Z}^*/n$.

# 3 CQE Test as a Rank Score Test

Given the comprehensive linear model that involves both $\boldsymbol{x}_i$ and $\boldsymbol{z}_i^*$:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{z}_i^{*\prime}\boldsymbol{\gamma}^* + e_i,$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}^*$ are parameter vectors, and $e_i$ is the error term, consider the null hypothesis $H_0 : \boldsymbol{\gamma}^* = \boldsymbol{0}$. Let $\hat{a}_i(t)$ denote the rank generating function based on the residual of the constrained model: $y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i$. That is, $\hat{a}_i(t) = 1$ if $y_i > \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(t)$; $\hat{a}_i(t) = 0$ if $y_i < \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(t)$; otherwise, $\hat{a}_i(t)$ is between zero and one. For a score-generating function $\varphi$, define $\hat{\boldsymbol{b}}_n(\varphi)$ as the vector of rank scores with the $i$-th element:

$$\hat{b}_i(\varphi) = -\int_0^1 \varphi(t)\,\mathrm{d}\hat{a}_i(t), \quad i = 1, \ldots, n,$$

and $A^2(\varphi) = \int_0^1 \left[\varphi(t) - \int_0^1 \varphi(t)\,\mathrm{d}t\right]^2 \mathrm{d}t$. To test $\boldsymbol{\gamma}^* = \boldsymbol{0}$, the regression rank score test of Gutenbrunner *et al.* (1993) is:

$$\mathrm{RS}_n = \boldsymbol{r}_n'(\varphi)\widehat{\boldsymbol{\Sigma}}_n^{-1}\boldsymbol{r}_n(\varphi)/A^2(\varphi) \xrightarrow{D} \chi^2(q-r), \tag{6}$$

where $\boldsymbol{r}_n(\varphi) = n^{-1/2}(\boldsymbol{Z}^* - \widehat{\boldsymbol{Z}}^*)'\hat{\boldsymbol{b}}_n(\varphi)$, $\widehat{\boldsymbol{\Sigma}}_n = (\boldsymbol{Z}^* - \widehat{\boldsymbol{Z}}^*)'(\boldsymbol{Z}^* - \widehat{\boldsymbol{Z}}^*)/n$, and $\widehat{\boldsymbol{Z}}^* = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}^*$ is the least-squares projection of $\boldsymbol{Z}^*$ on $\boldsymbol{X}$, with the $i$-th row $\hat{\boldsymbol{z}}_i^{*\prime} = \boldsymbol{x}_i'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}^*$.

To focus on the $\tau$-th quantile, we set $\varphi$ as $\varphi_\tau(t) = \tau - \mathbf{1}(t < \tau)$, the $\tau$-quantile score function. It is easy to verify that $A^2(\varphi_\tau) = \tau(1-\tau)$ and $\hat{b}_i(\varphi_\tau) = \tau - [1 - \hat{a}_i(\tau)]$. By invoking the definition of $\hat{a}_i(\tau)$,

$$\begin{aligned}\boldsymbol{r}_n(\varphi_\tau) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n (\boldsymbol{z}_i^* - \hat{\boldsymbol{z}}_i^*)[\tau - \mathbf{1}(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0)] + o_{\mathbb{P}}(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{z}_i^*[\tau - \mathbf{1}(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0)] + o_{\mathbb{P}}(1),\end{aligned}$$

where the last equality follows because $n^{-1/2}\sum_{i=1}^n \hat{\boldsymbol{z}}_i^*[\tau - \mathbf{1}(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0)]$ is $o_{\mathbb{P}}(1)$ by the asymptotic first-order condition. Thus, $\boldsymbol{r}_n(\varphi_\tau)$ is asymptotically equivalent to $\hat{\boldsymbol{\psi}}_n^*(\hat{\boldsymbol{\beta}}_n(\tau))$ in (3) and $\widehat{\boldsymbol{\Sigma}}_n = \boldsymbol{Z}^{*\prime}[\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{Z}^*/n$ is $\widehat{\boldsymbol{V}}_{n,\tau}$ under conditional homogeneity. This shows asymptotic equivalence between the CQE test and the rank score test (6). This result is consistent with the argument in Bontemps and Mizon (2008) that there is a comprehensive model underlying an encompassing test.

Table 1: Empirical sizes of the proposed test: Nominal size 5%.

| | $\tau = 0.25$ | | | | $\tau = 0.5$ | | | | $\tau = 0.75$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $q=2$ | $q=4$ | $q=6$ | $q=8$ | $q=2$ | $q=4$ | $q=6$ | $q=8$ | $q=2$ | $q=4$ | $q=6$ | $q=8$ |
| $p=2$ | 5.0 | 5.5 | 4.4 | 4.8 | 4.7 | 5.4 | 5.6 | 4.5 | 5.0 | 4.7 | 4.2 | 4.9 |
| $p=4$ | 5.4 | 5.0 | 5.1 | 5.5 | 5.5 | 5.8 | 5.5 | 5.8 | 5.6 | 5.1 | 5.4 | 4.8 |
| $p=6$ | 6.1 | 5.5 | 5.7 | 6.6 | 5.2 | 6.2 | 5.7 | 6.3 | 5.7 | 5.7 | 5.8 | 5.8 |
| $p=8$ | 6.3 | 6.6 | 6.9 | 6.2 | 6.1 | 5.7 | 6.5 | 6.8 | 6.5 | 6.4 | 6.5 | 6.6 |
| $p=2$ | 5.5 | 4.9 | 5.5 | 5.0 | 4.6 | 5.0 | 5.3 | 5.6 | 4.5 | 5.0 | 5.7 | 5.3 |
| $p=4$ | 4.6 | 5.6 | 4.6 | 4.9 | 5.2 | 5.3 | 5.1 | 5.4 | 5.4 | 5.2 | 5.3 | 4.5 |
| $p=6$ | 4.9 | 5.2 | 6.0 | 5.3 | 5.9 | 5.4 | 5.5 | 6.0 | 4.9 | 5.5 | 4.6 | 5.8 |
| $p=8$ | 5.5 | 6.6 | 5.5 | 5.0 | 5.5 | 5.3 | 4.6 | 5.0 | 6.0 | 5.2 | 5.1 | 4.9 |

Note: Upper and lower panels are the results for $n = 100$ and $n = 300$, respectively; all entries are in percentages.

# 4 Monte Carlo Simulations

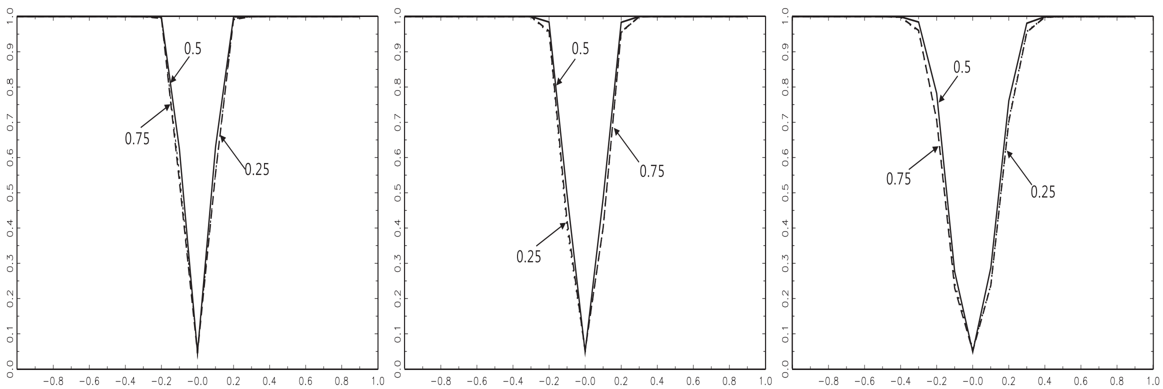
In our simulations, we consider the following data generating process (DGP):

$$y_i = (1-\lambda)\boldsymbol{x}_i'\boldsymbol{\beta} + \lambda\boldsymbol{z}_i'\boldsymbol{\gamma} + \varepsilon_i, \quad i = 1, \ldots, n, \ \lambda \in [-1, 1], \tag{7}$$

where $\boldsymbol{x}_i$ $(p \times 1)$ and $\boldsymbol{z}_i$ $(q \times 1)$ both contain a constant term, $\varepsilon_i$ and the remaining elements of $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are i.i.d. $\mathcal{N}(0, 1)$. Both $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are set to vector of ones. For the $\tau$-th QR, the error term $\varepsilon_i$ satisfies the restriction that its $\tau$-th conditional quantile is zero. For example, when $\tau = 0.25, 0.5, 0.75$, the errors are $\mathcal{N}(0, 1)$ plus $0.6475, 0, -0.6475$, respectively. We consider different numbers of regressors ($p, q = 2, 3, \ldots, 8$) and different samples ($n = 100, 300$); the number of replications is 3000.

Note that the null model is (7) with $\lambda = 0$, and the alternative model is (7) with $\lambda = 1$. Table 1 contains the empirical sizes of the proposed test for $\tau = 0.25, 0.5, 0.75$ and the nominal size 5%; the results for other $\tau$ and nominal sizes are qualitatively similar. To conserve space, we report only the results for $p, q = 2, 4, 6, 8$. It can be seen that, when $n = 100$, the proposed test is properly sized when $p$ and $q$ are not too large; otherwise it may be slightly over-sized. Yet, the size distortion becomes smaller when the sample increases.

In the power simulations, we consider the alternative model with $(p, q) = (2, 6), (4, 4), (6, 2)$ and $\tau = 0.25, 0.5, 0.75$. We plot the power functions against $\lambda$ for the sample $n = 300$ in Figures 1. It can be seen that, for given $(p, q)$, the power functions for different $\tau$ are close to each other. In particular, each power function increases with the magnitude of $\lambda$ and reaches power one quickly, and it is symmetric about $\lambda = 0$. For a given $\lambda$ that is not too big, the




$(p,q) = (2,6)$ $(p,q) = (4,4)$ $(p,q) = (6,2)$

Figure 1: Empirical power functions against $\lambda$ values: $n = 300$.

power is higher when $q$ (the number of regressors in the alternative model) is larger. The power functions for $n = 100$ (not reported) are of similar shapes but with smaller values.

# 5 Conclusions

In this paper we propose a CQE test for linear QR models. This test is an extension of the CME test to QR models and also a regression rank score test. Our simulations confirm that this test has good finite sample performance.

## Appendix

**Proof of Equation (4)**: We first write

$$\begin{aligned}
\hat{\psi}_n^*(\hat{\boldsymbol{\beta}}_n(\tau)) = {} & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{z}_i^* \big[\tau - \mathbf{1}\big(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0\big)\big] \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{z}_i^* \big[\mathbf{1}\big(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}(\tau) < 0\big) - F_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big)\big] \\
& - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{z}_i^* \big[\mathbf{1}\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau) < 0\big) - F_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau)\big)\big] \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{z}_i^* \big[F_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big) - F_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau)\big)\big].
\end{aligned}$$

By Lemma 1 of Gutenbrunner and Jurečková (1992) and Theorem 3.3 of Gutenbrunner *et al.* (1993), the second and third terms on the right-hand side vanish in probability, uniformly in $\tau$. For the fourth term, the Taylor expansion of yields

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{z}_i^*\left(F_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big)-F_{y_i|\mathcal{F}}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n(\tau))\right)\\
&\quad=-\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{z}_i^* f_{y_i|\mathcal{F}_i}\big(\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)\big)\boldsymbol{x}_i'\right)\left(\sqrt{n}\big[\hat{\boldsymbol{\beta}}_n(\tau)-\boldsymbol{\beta}(\tau)\big]\right)+o_{\mathbb{P}}(1)\\
&\quad=-\boldsymbol{M}_{fzx}\boldsymbol{M}_{fxx}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{x}_i\big[\tau-\mathbf{1}\big(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)<0\big)\big]+o_{\mathbb{P}}(1),
\end{aligned}$$

where the last equality follows from the Bahadur representation. It follows that

$$\hat{\psi}_n^*(\hat{\boldsymbol{\beta}}_n(\tau))=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\boldsymbol{z}_i^*-\boldsymbol{M}_{fzx}\boldsymbol{M}_{fxx}^{-1}\boldsymbol{x}_i\big)\big[\tau-\mathbf{1}\big(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}(\tau)<0\big)\big]+o_{\mathbb{P}}(1). \quad \square$$

## References


Atkinson, A. (1970). A method for discriminating between models, *Journal of the Royal Statistical Society, Series B*, **32**, 323–353.

Bontemps, C. and G. Mizon (2008). Encompassing: Concepts and Implementation, *Oxford Bulletin of Economics and Statistics*, **70**, 721–750.

Chen, Y.-T. and C.-M. Kuan (2002). The pseudo-true score encompassing test for non-nested hypotheses, *Journal of Econometrics*, **106**, 271–295.

Chen, Y.-T. and C.-M. Kuan (2007). Corrigendum to "The pseudo-true score encompassing test for non-nested hypotheses," *Journal of Econometrics*, **141**, 1412–1417.

Davidson, R. and J. MacKinnon (1981). Several tests for model specification in the presence of alternative hypotheses, *Econometrica*, **49**, 781–793.

Fisher, G. and M. McAleer (1981). Alternative procedures and associated tests of significance for non-nested hypotheses, *Journal of Econometrics*, **16**, 103–119.

Gouriéroux, C., A. Monfort, and A. Trognon (1983). Testing nested or non-nested hypotheses, *Journal of Econometrics*, **21**, 83–115.

Gutenbrunner, C. and J. Jurečkova (1992). Regression rank scores and regression quantiles, *The Annals of Statistics*, **20**, 305–330.

Gutenbrunner, C., J. Jurečkova, R. Koenker, and S. Portnoy (1993). Tests of linear hypotheses based on regression rank scores, *Nonparametric Statistics*, **2**, 307–331.

Koenker, R. (2005). *Quantile Regression*, New York: Cambridge University Press.

Koenker, R. and G. Bassett (1978). Regression quantile, *Econometrica*, **46**, 33–50.

MacKinnon, J., H. White, and R. Davidson (1983). Tests for model specification in the presence of alternative hypotheses, *Journal of Econometrics*, **21**, 53–70.

Mizon, G. (1984). The encompassing approach in econometrics, in: D. F. Hendry and K. F. Wallis (eds.), *Econometrics and Quantitative Economics*, 135–172, Oxford: Blackwell.

Mizon, G. and J. Richard (1986). The encompassing principle and its application to testing non-nested hypotheses, *Econometrica*, **54**, 657–678.

Otsu, T. (2008). Conditional empirical likelihood estimation and inference for quantile regression models, *Journal of Econometrics*, **142**, 508–538.

Powell, J. (1991). Estimation of monotonic regression models under quantile restrictions, in W. Barnett, J. Powell, and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics*, 357–384, Cambridge: Cambridge University Press.

Ramalho, J. and E. Smith (2002). Generalized empirical likelihood non-nested tests, *Journal of Econometrics*, **107**, 99–125.

Santos Silva, J. (2001). A score test for non-nested hypotheses with applications to discrete data models, *Journal of Applied Econometrics*, **16**, 577–597.

Smith, R. (1992). Non-nested tests for competing models estimated by generalized method of moments, *Econometrica*, **60**, 973–980.

Wooldridge, J. (1990). An encompassing approach to conditional mean tests with applications to testing nonnested hypotheses, *Journal of Econometrics*, **45**, 331–350.

White, H. (1999). *Asymptotic Theory for Econometricians*, Academic Press: San Diego.