Closed-based testing when multiple quantile regressions are fitted

Published Version:
De Santis, R., Vesely, A., Andreella, A. (2025). Closed-based testing when multiple quantile regressions are fitted. Cham : Springer [10.1007/978-3-031-95995-0_21].

Availability:
This version is available at: https://hdl.handle.net/11585/1021785 since: 2025-08-25

Published:
DOI: http://doi.org/10.1007/978-3-031-95995-0_21

(Article begins on next page)

Closed-based testing when multiple quantile regressions are fitted

De Santis Riccardo[1], Vesely Anna[2], and Andreella Angela[3]

[1] University of Padova, Department of Statistical Sciences, Padova, Italy,
riccardo.desantis@unipd.it
[2] University of Bologna, Department of Statistical Sciences, Bologna, Italy
[3] University of Trento, Department of Economics and Management, Trento, Italy

Abstract. This paper addresses the challenge of conducting multiple quantile regressions at different levels and the consequent issue of controlling the familywise error rate (FWER). Current practices in various fields typically involve conducting separate tests for each quantile, leading to a multiplicity problem that often remains unaddressed. We propose a method that integrates the Wald test within a closed-testing procedure to manage multiple tests effectively. We conduct simulation studies across various scenarios to demonstrate the efficacy of our method in controlling the FWER and its power compared to traditional approaches like the Bonferroni correction. Our findings advocate for a more rigorous application of statistical tests in quantile regressions to prevent false discoveries and enhance the reliability of analytical conclusions.

Keywords: quantile regression, type I error, closed testing, Wald test

1 Introduction

Quantile regression extends linear regression by estimating a conditional quantile of a response of interest, providing a more comprehensive view of the relationship between covariates and response beyond just the conditional mean. Fitting multiple quantile regressions at different quantile levels is crucial in many areas where the impact of some covariates varies across the distribution. For example, in finance, investment firms study the impact of macroeconomic variables on the lower quantiles of stock returns during economic downturns, focusing on worst-case scenarios for risk management and loss mitigation. In medical research, the effect of a treatment intended to reduce cholesterol may be more pronounced at the upper quantiles of the cholesterol distribution, particularly for patients with extremely high initial levels. Quantile regressions help identify which patients are most likely to benefit from the treatment, enhancing the effectiveness of medical interventions. In such contexts, analyzing a single quantile value alone could overlook significant results of interest.

We focus on the problem of testing the evidence of a covariate effect across multiple quantile levels within a set T. In general, the literature refers to *regional quantile testing* if $T \subset (0, 1)$ is a continuous range, and *global/partial quantile* *regression* if $T = \{\tau_1, \dots, \tau_k\}$ is a finite set, with global quantile regression analyzing a broad spectrum of quantiles and partial quantile regression focusing on a limited subset. Various methods exist for regional quantile testing, with examples found in the work of Sun et al. [17] and He et al. [9]. In contrast, a valid approach for conducting quantile regressions at multiple discrete k levels is lacking. Typically, researchers perform k separate regressions, analyzing each regression's results and the corresponding p-values/test statistics independently.

This practice, however, leads to a multiple testing problem, where performing k separate tests increases the risk of false discoveries if p-values are not adjusted to control family errors such as familywise error (FWER) or false discovery rates (FDR). The multiplicity problem is not commonly addressed within the quantile regression literature. To the best of our knowledge, only Mrkvicka et al. [16] consider multiple testing in this framework; yet, their proposal lacks a solid theoretical foundation.

The strategy of applying quantile regression across various quantiles and making inferences about the effects of covariates without addressing multiple testing issues is a common oversight in various disciplines. Numerous examples can be found in studies spanning markets [18], biology [1], management [3, 12], economics [8, 7], public health [11], health [5], education [2], energy [6], and climate [4]. This widespread practice highlights a methodological gap, emphasizing the need for the field to develop and apply rigorous statistical methods that properly account for multiple comparisons and ensure the validity of findings.

We propose a straightforward solution for handling multiple testing in cases where global or partial quantile regressions are of interest. We employ a Wald-type test [14], integrating it within a closed-testing procedure [15] to obtain strong control of the FWER. This approach enhances the flexibility of the analyses and ensures valid control of type I errors across multiple quantile levels, thereby improving the reliability of the statistical inferences made from quantile regression analyses.

The manuscript is organized as follows. Section 2 details the application of the Wald test as local test within the closed testing procedure we propose. Section 3 presents simulations that demonstrate the consequences of using uncorrected p-values, which fail to control the type I error rate, and illustrates how our proposed approach effectively manages the FWER while maintaining a competitive power compared to traditional methods like the Bonferroni correction. Finally, Section 4 is devoted to the conclusions, summarizing the key findings and implications of our research.

2 Methodology

Consider a response variable vector $\mathbf{y} = (y_1, \dots, y_n)^\top$, with entries that are not necessarily independent and identically distributed (i.i.d.), and two continuous covariates: a covariate of interest $\mathbf{x} = (x_1, \dots, x_n)^\top$ and a nuisance covariate $\mathbf{z} = (z_1, \dots, z_n)^\top$. For a given level $\tau \in (0, 1)$, we assume that the conditional

τ-th quantile of $\mathbf{y}$, given $\mathbf{x}$ and $\mathbf{z}$, is

$$Q_{y_i|x_i,z_i}(\tau \mid x_i, z_i) = \alpha + \beta(\tau)x_i + \gamma(\tau)z_i$$

where $\beta(\tau)$ is the quantile-specific coefficient for $\mathbf{x}$, and $\gamma(\tau)$ for $\mathbf{z}$. This setting can be directly extended to more general scenarios.

The aim is to test the effect of the target covariate $\mathbf{x}$ at k different quantile levels $\tau_1, \dots, \tau_k$. Accordingly, we formulate the following set of k null hypotheses, each tested against a two-sided alternative:

$$\begin{cases} H_0^1 : \beta(\tau_1) = 0 \\ \vdots \\ H_0^k : \beta(\tau_k) = 0 \end{cases}$$

To test these null hypotheses with strong control of the FWER at a desired level $\alpha \in (0, 1)$, we employ the closed testing procedure [15]. This method rejects any individual H_0^m if all possible intersection hypotheses involving it, i.e., all

$$H_V = \bigcap_{v \in V} H_0^v \quad \text{where} \quad m \in V \subseteq \{1, \dots, k\},$$

can be rejected using a valid local α-level test.

As local test, we use the Wald-type test proposed by Koenker and Portnoy (1996) [14], particularly applicable in the presence of non-i.i.d. errors; see [14] for specific assumptions. The test statistic in the case where $V = \{1, \dots, k\}$ is the full set is

$$\sqrt{n}\,\hat{\boldsymbol{\beta}}^\top \hat{\Sigma}\,\hat{\boldsymbol{\beta}} \dot{\sim} \chi_k^2 \tag{1}$$

and analogous definitions apply to any subset. Here $\hat{\boldsymbol{\beta}} = (\hat{\beta}(\tau_1), \dots, \hat{\beta}(\tau_k))^\top$ is the estimated vector of coefficients and $\hat{\Sigma}$ is the estimated asymptotic covariance matrix. In presence of the intercept, $\mathbf{x}$ and $\mathbf{z}$ (where the first, $\mathbf{x}$, is of interest), the generic element is $\hat{\Sigma}_{i,j} = A_{2,2}$, with

$$A = P(\tau_i)^{-1} J(\tau_i, \tau_j) P(\tau_j)^{-1} \in \mathbb{R}^{3\times 3}.$$

$P(\tau_i), P(\tau_j) \in \mathbb{R}^{3\times 3}$ are estimated using the sparsity estimation method proposed by Hendricks and Koenker (1992) [10], while

$$J(\tau_i, \tau_j) = [\tau_i \wedge \tau_j - \tau_i\tau_j] n^{-1} \sum_{h=1}^{n} (1, x_h, z_h)^\top (1, x_h, z_h) \in \mathbb{R}^{3\times 3},$$

where $\wedge$ denotes the minimum between the two quantile levels.

Although the test proposed by Koenker in his seminal work on quantile regression [14] is foundational, it is not implemented in the widely used `R` package for quantile regression, `quantreg` [13]. However, in scenarios where $k = 1$, the test defined in Equation (1) aligns with the approach utilized by the `quantreg` package when errors are not independent and identically distributed; this is specified using the `se = "nid"` option in the `summary` function of the package.

3 Simulation study

A simulation study is conducted to evaluate the control of type I error and the power of both the Bonferroni correction and the proposed Wald-based closed-testing approach, highlighting the importance of these corrections for statistical validity. The study involves 1000 replicates, each with a sample size $n = 1000$. The two covariates $\mathbf{x}$ and $\mathbf{z}$ are drawn from a multivariate normal distribution; with the exception of the first examined scenario, they have means 1 and 0, standard deviations 2.5 and 1, and varying correlations ρ. The response $\mathbf{y}$ is generated in different ways, where each setting is designed to test different statistical behaviors. The analysis considers five quantiles, specifically $T = \{0.1, 0.25, 0.5, 0.75, 0.9\}$, testing the effect of $\mathbf{x}$ across the distribution at level $\alpha = 0.05$.

It is important to note that both methods — the Bonferroni correction and the Wald-based closed-testing approach — presume that the univariate tests adequately control the type I error. This control is theoretically established as an asymptotic result. However, despite the large sample size used in our simulations, we observe a slightly anti-conservative behavior at the extreme quantiles. Although these findings are pertinent, they fall outside the primary focus of this paper. Therefore, we will not detail these results here and will proceed under the assumption that the univariate tests are valid for the purpose of our analysis.

First, we evaluate the control of the FWER in four scenarios where $\mathbf{x}$ has no effect on $\mathbf{y}$:

1. x_i, z_i and y_i follow independent standard normal distributions $\mathcal{N}(0, 1)$.
2. $\rho = 0.3$ and $y_i \sim \mathcal{N}(0, 1)$.
3. $\rho = 0$ and $y_i \sim \mathcal{N}(1, \sigma_i)$, where σ_i is sampled from a uniform distribution $\mathcal{U}[0.5, 3]$.
4. $\rho = 0.5$ and y_i follows a Student's t-distribution with degrees of freedom $\nu_i \sim \mathcal{U}\{5, \ldots, 10\}$.

Results (Table 1) distinctly show that failing to correct for multiplicity, which is commonly observed as per our discussion in the introduction, leads to a significantly inflated FWER. In contrast, both the Bonferroni and the closed testing approaches effectively control the FWER. This demonstrates the critical importance of applying multiplicity corrections when multiple quantile regressions are fitted to avoid erroneous conclusions.

Subsequently, we compare the power of the Bonferroni method and our proposal in the following scenarios.

5. $\rho = 0$ and $y_i \sim \mathcal{N}(0.1 \cdot x_i, 1)$.
6. $\rho = 0.5$ and $y_i \sim \mathcal{N}(1, \sigma_i)$, where $\sigma_i = s_i + 0.2 \cdot |x_i|$ and $s_i = \mathcal{U}[0.5, 2]$.
7. $\rho = 0.5$ and $y_i \sim \mathcal{N}(1 + 0.1x_i, \sigma_i)$, where $\sigma_i = s_i + 0.1 \cdot |x_i|$.
8. $\rho = 0.5$ and y_i follows a noncentral t-distribution with degrees of freedom $\nu_i \sim \mathcal{U}\{5, \ldots, 10\}$ and noncentrality parameter $\mu_i = 1 + 0.1 \cdot x_1$.

Results are shown in Table 2, where in most scenarios the proposed approach gains more power compared to the traditional Bonferroni method.

Table 1. Empirical FWER under four scenarios using uncorrected p-values, p-values corrected by Bonferroni and the proposed approach, i.e., closed testing (CT). The (local) tests used are the Wald tests defined in Equation (1)

	Uncorrected	**Bonferroni**	**Wald CT**
Scenario 1	0.206	0.054	0.028
Scenario 2	0.215	0.056	0.019
Scenario 3	0.192	0.057	0.026
Scenario 4	0.213	0.054	0.022

4 Conclusions

Applying multiple quantile regressions across different quantile levels without accounting for multiple comparisons is a common practice in several fields. This study highlights the need for rigorous multiple testing adjustments in this framework. By integrating the Wald test within a closed-testing framework, our approach effectively controls the FWER, enhancing the reliability of statistical inferences. Simulations show that uncorrected tests lead to inflated type I error rates, while the proposed methodology not only maintains appropriate error contol but also shows greater power in detecting true effects compared to traditional methods like the Bonferroni correction. This advancement can significantly improve the credibility and accuracy of findings in quantitative research.

References

1. Agarwal, G., Saade, S., Shahid, M., Tester, M., Sun, Y.: Quantile function modeling with application to salinity tolerance analysis of plant data. BMC Plant Biol. 19, 1–12 (2019)
2. Bassett, G.W.Jr, Tam, M.Y.S., Knight, K.: Quantile models and estimators for data analysis. Metrika 55, 17–26 (2002)
3. Bien, Y.H., Su, Y.M., Wang, Y.C., Chang, C.H., Ming-Ru, Z.: Study on dynamic factors striding across the middle income trap-transnational empirical analysis based on quartile regression. J. Stat. Manag. Syst. 19(1), 107–118 (2016)
4. Chamaillé-Jammes, S., Fritz, H., Murindagomo, F.: Detecting climate changes of concern in highly variable environments: quantile regressions reveal that droughts worsen in Hwange National Park, Zimbabwe. J. Arid Environ. 71(3), 321–326 (2007)
5. Cheah, Y.K., Goh, K.L., Abdul Adzis, A.: Malaysian household expenditure on tobacco: evidence from pooled cross-sectional quantile estimations. Drugs Habits Soc. Policy 23(1), 79–90 (2022)
6. Cheng, F., Yang, S., Zhou, K.: Quantile partial adjustment model with application to predicting energy demand in China. Energy 191, 116519 (2020)
7. Falk, M.: Quantile estimates of the impact of R&D intensity on firm performance. Small Bus. Econ. 39, 19–37 (2012)
8. Fousekis, P., Lazaridis, P.: The demand for selected nutrients by Greek households: an empirical analysis with quantile regressions. Agric. Econ. 32(3), 267–279 (2005)
9. He, X., Hsu, Y.H., Hu, M.: Detection of treatment effects by covariate-adjusted expected shortfall. Ann. Appl. Stat. 4, 2114–2125 (2010)

Table 2. Empirical power under four scenarios using p-values corrected by Bonferroni and the proposed approach, i.e., closed testing (CT). The (local) tests used are the Wald tests defined in Equation (1)

	Quantile	**Bonferroni**	**Wald CT**
Scenario 5	0.10	0.616	0.799
	0.25	0.861	0.913
	0.50	0.915	0.944
	0.75	0.861	0.910
	0.90	0.628	0.789
Scenario 6	0.10	0.223	0.227
	0.25	0.161	0.180
	0.50	0.025	0.046
	0.75	0.178	0.206
	0.90	0.220	0.226
Scenario 7	0.10	0.030	0.075
	0.25	0.208	0.274
	0.50	0.668	0.636
	0.75	0.777	0.730
	0.90	0.527	0.555
Scenario 8	0.10	0.569	0.712
	0.25	0.789	0.824
	0.50	0.845	0.856
	0.75	0.695	0.758
	0.90	0.407	0.581

10. Hendricks, W., Koenker, R.: Hierarchical spline models for conditional quantiles and the demand for electricity. J. Am. Stat. Assoc. 87(417), 58–68 (1992)
11. Juvanhol, L.L., Lana, R.M., Cabrelli, R., Bastos, L.S., Nobre, A.A., Rotenberg, L., Griep, R.H.: Factors associated with overweight: are the conclusions influenced by choice of the regression method? BMC Public Health 16, 1–9 (2016)
12. Karimi Alavijeh, N., Zangoei, S.: How R&D expenditure affects renewable energy development: the role of trade liberalization and CO2 emissions in G-7 countries. Int. J. Energy Sect. Manag. 18(6), 1263–1284 (2024)
13. Koenker, R.: quantreg: Quantile Regression. R package version 5.98 (2024)
14. Koenker, R., Portnoy, S.: Quantile regression. ABE (1996)
15. Marcus, R., Peritz, E., Gabriel, K.R.: On closed testing procedures with special reference to ordered analysis of variance. Biometrika 63(3), 655–660 (1976)
16. Mrkvička, T., Konstantinou, K., Kuronen, M., Myllymäki, M.: Global quantile regression. ArXiv preprint, arXiv:2309.04746 (2023)
17. Sun, Y., He, X.: Model-based bootstrap for detection of regional quantile treatment effects. J. Nonparametric Stat. 33(2), 299–320 (2021)
18. Wang, Y.H., Hung, J.C., Lee, Y.H., Chuang, C.C.: Computing regression quantiles to analysis the relationship between market behavior and political risk. Qual. Quant. 46, 1047–1055 (2012)