Bootstrapping Quasi Likelihood Ratio Tests under Misspecification

Patrice Bertail*
MODAL'X, UPL, Université Paris-Nanterre
and
Pascal Lavergne†
Toulouse School of Economics, Université Toulouse Capitole

May 11, 2020

Abstract

We consider quasi likelihood ratio (QLR) tests for restrictions on parameters under potential model misspecification. For convex M-estimation, including quantile regression, we propose a general and simple nonparametric bootstrap procedure that yields asymptotically valid critical values. The method modifies the bootstrap objective function to mimic what happens under the null hypothesis. When testing for an univariate restriction, we show how the test statistic can be made asymptotically pivotal. Our bootstrap can then provide asymptotic refinements as illustrated for a linear regression model. A Monte-Carlo study and an empirical application illustrate that double bootstrap of the QLR test controls level well and is powerful.

Keywords: Hypothesis Testing, Asymptotic Refinements.

---

*This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023- 01).

†Pascal Lavergne acknowledges support from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). Address correspondence: Pascal Lavergne, Toulouse School of Economics, Université Toulouse Capitole, 1 Esplanade de l'Université, 31000 Toulouse FRANCE. Email: pascal.lavergne@ut-capitole.fr

1 Introduction

Let $z \in R^k$ be a random vector and $Q(\theta) = \mathbb{E}\, q(z,\theta)$ a function from R^p to R that admits a unique minimizer $\theta_\dagger$ on the parameter space Θ, i.e.

$$\theta_\dagger = \arg\min_{\Theta} Q(\theta)\,. \tag{1.1}$$

For a random sample $\{z_i, i = 1, \dots n\}$ from z, a consistent M-estimator of $\theta_\dagger$ is defined as

$$\widehat{\theta}_n = \arg\min_{\Theta} Q_n(\theta)\,, \qquad Q_n(\theta) = n^{-1}\sum_{i=1}^{n} q(z_i,\theta)\,.$$

We study tests on parameters based on quasi likelihood ratio (QLR) statistics. We consider for simplicity linear restrictions on parameters of the type

$$H_0:\ d'\,(\theta_\dagger - h) = \mathbf{0}, \quad \text{against} \quad H_A:\ d'\,(\theta_\dagger - h) \neq \mathbf{0}\,,$$

where d is a known $p \times r$ full rank matrix, $r < p$ and h is a known vector in R^p. Let

$$\theta_\dagger^0 = \arg\min_{\Theta\cap H_0} Q(\theta)\,. \tag{1.2}$$

The estimator subject to the null constraint is $\widehat{\theta}_n^0 = \arg\min_{\Theta\cap H_0} Q_n(\theta)$. The Quasi Likelihood Ratio (QLR) statistic for testing H_0 is

$$\mathrm{QLR}_n = 2\,n\,\left[Q_n(\widehat{\theta}_n^0) - Q_n(\widehat{\theta}_n)\right]\,.$$

The main advantages of the QLR statistic are (i) it avoids estimation of the asymptotic covariance matrix of estimators, by contrast to Wald and Score statistics (ii) it is transformation equivariant. However, under misspecification, i.e. outside model assumptions, it is generally not asymptotically pivotal under H_0 due to the potential failure of Bartlett's second identity, see Foutz and Srivastava (1977) and Kent (1982) for unconditional Maximum Likelihood (ML), Vuong (1989) for conditional ML, and Marcellino and Rossi (2008) for M-estimation. We revisit these results for a convex but potentially non-differentiable M-estimation criterion, thus allowing for quantile regression. When testing for a single restriction, a robust QLR statistic, which is asymptotically chi-square even under misspecification, easily obtains, as in Stafford (1996). The construction is similar to robust Score and Wald statistics, see e.g. White (1982) and Royall (1986).

Our main focus is to develop a consistent bootstrap procedure based on nonparametric naive bootstrap resampling. To make the bootstrap statistic behaves as if the null hypothesis were true, we modify the criterion function to be optimized. We label our method *bootstrapping under the null hypothesis.* It is shown to be consistent under misspecification. Moreover, it yields asymptotic higher-order refinements in the regression model. Our theoretical results and our simulation study show that bootstrapping under the null controls size better than competitors under misspecification. Our simulations also compare the small sample performance of two competing procedures when testing a single restriction, namely bootstrapping the robust QLR statistic or double bootstrapping the non-robust QLR statistic, see Beran (1988). Our findings suggest that the first procedure may be preferable in moderate samples, because it avoids estimation of Hessian and score variance matrices, in line with the conclusions of Stafford (1996), who states "the use of a model-robust variance estimate for the signed square root, score or Wald statistic, while leaving bias and skewness characteristics relatively unchanged, can increase variability considerably."

The study of potentially misspecified models dates back at least to Huber (1967) and Eicker (1967). For conditional ML, Gourieroux, Monfort, and Trognon (1984) have shown that the parameters of the conditional mean can be estimated consistently whenever it is well specified even if the likelihood itself is not. Similarly, ordinary least-squares consistently estimate mean parameters in the presence of heteroscedasticity of unknown form. If the true conditional expectation is nonlinear, OLS provides the best linear approximation to the true conditional expectation that can be of interest on its own, see White (1980). Building on Eicker (1967), White (1980, 1982) and Royall (1986) have shown how Wald and score statistics can be rendered robust to misspecification. However their small sample behavior are sensitive to the implementation details and the design of the data, see Mackinnon and White (1985) and Chesher and Austin (1991) among others. Bootstrap methods have been investigated to obtain better approximations, see e.g. Wu (1986), Mammen (1992), Aerts and Claeskens (2001), Andrews (2002), and Gonçalves and White (2004). Kline and Santos (2012) recently pointed out that in a regression model the popular wild bootstrap does not improve over the asymptotic approximation for the Wald statistic when the model is misspecified. By contrast, the naive bootstrap of Efron (1979) provides asymptotic improvements for the robust Wald test as shown by Hall and Horowitz (1996).

Recent work has focused on bootstrap methods for QLR tests. The most advocated method computes a bootstrap statistic that does not test H_0 but rather $d'\left(\theta_{\dagger}-h\right) = d'\left(\widehat{\theta}_n - h\right)$, where $\widehat{\theta}_n$ is the initial estimator (considered as fixed for the bootstrap distribution). Kim (2003) propose a parametric bootstrap and establishes higher-order improvements for a well specified model. Camponovo (2016) establishes higher-order improvements of a block bootstrap test in a dependent data context assuming the equivalent of second Bartlett's identity. Spokoiny and Zhilova (2015) study the multiplier bootstrap, allowing for dependence and a large number of parameters, and show that it may not be valid for some misspecified models, see Chen and Pouzo (2009) for a similar method in a semiparametric context. For quantile regression, Angrist, Chernozhukov, and Fernández-Val (2006) rely on subsampling. Lee and Yang (2020) study the m out of n bootstraps of QLR tests from M-estimation.

In Section 2, we present our bootstrap method and we detail implementation for linear regression and linear quantile regression. In Section 3, we focus on asymptotic properties of QLR statistics under misspecification and the construction of robust statistics, allowing for a non-differentiable criterion. We then show consistency of bootstrapping under the null and we establish asymptotic higher-order refinements of the bootstrap robust QLR test in the linear regression model. Section 4 gathers small sample evidence on the behavior of our method compared to existing methods. We then report some empirical results for quantile regression of children birthweight. Section 5 concludes. Section 6 gathers our technical proofs.

2 Bootstrapping Under the Null

We assume a convex criterion that is well approximated by a quadratic function, that is

$$Q_n(\theta+t)-Q_n(\theta)=n^{-1}\sum_{i=1}^{n}D'(z_i,\theta)t+\frac{1}{2}t'A(\theta)t+R_n(t)\,,\quad R_n(t)=o_p(\|t\|^2),\ \forall\theta,t\,. \tag{2.3}$$

Let the empirical score function be the derivative of the quadratic approximation of $Q_n(\cdot)$

$$S_n(\theta)=n^{-1}\sum_{i=1}^{n}D(z_i,\theta)\,.$$

If $Q_n(\cdot)$ is differentiable, this is simply its derivative. To bootstrap the QLR statistic, let $Q_n^*(\theta)$ be the criterion based on observations $\{z_i^*, i=1,\ldots n\}$ resampled with replacement from the original data. We then consider the modified criterion

$$\widehat{Q}_n^*(\theta)=Q_n^*(\theta)-S_n'(\widehat{\theta}_n^0)(\theta-\widehat{\theta}_n^0)\,. \tag{2.4}$$

The bootstrap and constrained bootstrap estimator are defined as

$$\widehat{\theta}_n^* = \arg\min_{\Theta} \widehat{Q}_n^*(\theta) \quad \text{and} \quad \widehat{\theta}_n^{0*} = \arg\min_{\Theta\cap H_0} \widehat{Q}_n^*(\theta)\,.$$

The bootstrap QLR statistic is

$$\text{QLR}_n^* = 2\,n\,\left[\widehat{Q}_n^*(\widehat{\theta}_n^{0*}) - \widehat{Q}_n^*(\widehat{\theta}_n^*)\right] = 2\,n\,\left[Q_n^*(\widehat{\theta}_n^{0*}) - Q_n^*(\widehat{\theta}_n^*) - S_n'(\widehat{\theta}_n^0)(\widehat{\theta}_n^{0*} - \widehat{\theta}_n^*)\right]\,.$$

We label our method *bootstrapping under the null* because it makes the bootstrap estimator $\widehat{\theta}_n^*$ and QLR behave as if the restrictions were true. A useful intuition can help to understand why. Since $Q_n^*(\theta)$ converges to $Q_n(\theta)$ conditionally on the initial sample, at the limit (2.4) becomes $Q_n(\theta) - S_n'(\widehat{\theta}_n^0)\left(\theta - \widehat{\theta}_n^0\right)$. But the constrained estimator $\widehat{\theta}_n^0$ under H_0 is obtained by solving the Lagrangian program $\min_{\theta,\mu} Q_n(\theta) - \mu' d' (\theta - h)$, where at the optimum $d\,\mu$ is the empirical score at $\widehat{\theta}_n^0$. So, conditionally on the initial sample, $\widehat{\theta}_n^*$ becomes close to $\widehat{\theta}_n^0$. Since clearly $\widehat{\theta}_n^{0*}$ also becomes close to $\widehat{\theta}_n^0$, the bootstrapped statistic is bounded in probability irrespective to whether H_0 holds and has an asymptotic distribution which mimics the one of QLR_n under H_0.

Our method equates the derivatives of the quadratic approximation of $Q_n^*(\theta)$ at $\widehat{\theta}_n^*$ to $S_n(\widehat{\theta}_n^0)$. Hall and Horowitz (1996) and Andrews (2002) study related “recentering” methods for constructing Wald statistics. The formers use it to account for the non-nullity of the empirical moments in overidentified models estimated by Generalized Method of Moments, the latter because the average score evaluated using block-bootstrapped data can be different from zero. Here we use it instead to account for the fact that the score evaluated at the constrained estimator is not zero.

Least-Squares Regression. Consider the linear model

$$y_i = x_i'\theta + \varepsilon_i \qquad i = 1, ..., n$$

where $z_i = (y_i, x_i), i = 1, ..., n$ are independent and identically distributed from P. We allow for conditional heteroscedasticity in the error term ε and for correlation between ε and x. The unique minimizer of $Q(\theta) = \mathbb{E}\,(y - x'\theta)^2/2$ is $\theta_\dagger = \mathbb{E}^{\,-1}(xx')\mathbb{E}\,(x'y)$, where $\mathbb{E}$ denotes expectation with respect to P. The corresponding least-squares estimator is $\widehat{\theta}_n = \mathbb{E}_n^{\,-1}(xx')\mathbb{E}_n(x'y)$, where $\mathbb{E}_n$ denotes expectation with respect to the empirical distribution of the data P_n. Similarly, the constrained minimizer of $Q(\theta)$ is

$$\theta_\dagger^0 = \theta_\dagger - \mathbb{E}^{\,-1}(xx')d\left(d'\mathbb{E}^{\,-1}(xx')d\right)^{-1} d'\left(\theta_\dagger - h\right) \,,$$

and the constrained estimator is

$$\widehat{\theta}_n^0 = \widehat{\theta}_n - \mathbb{E}_n^{\,-1}(xx')d\left(d'\mathbb{E}_n^{\,-1}(xx')d\right)^{-1} d'\left(\widehat{\theta}_n - h\right) \,.$$

Moreover, as $Q_n(\theta)$ is quadratic in θ and $\widehat{\theta}_n$ minimizes $Q_n(\theta)$,

$$\begin{aligned} QLR_n &= 2\,n\left[Q_n(\widehat{\theta}_n^0) - Q_n(\widehat{\theta}_n)\right] = n\,\left(\widehat{\theta}_n - \widehat{\theta}_n^0\right)\mathbb{E}_n(xx')\left(\widehat{\theta}_n - \widehat{\theta}_n^0\right)' \\ &= \left(d'\left(\widehat{\theta}_n - h\right)\right)'\left(d'\mathbb{E}_n^{\,-1}(xx)d\right)^{-1} d'\left(\widehat{\theta}_n - h\right) \,, \end{aligned}$$

and is thus equal to a Wald or Score statistic for testing H_0 up to the residual variance $\widehat{\sigma}^2$.

In the bootstrap world, the modified criterion makes $\widehat{\theta}_n^*$ converging to $\widehat{\theta}_n^0$ as

$$\widehat{\theta}_n^* = \mathbb{E}_n^{*^{-1}}(xx')\left(\mathbb{E}_n^*(xy) + S_n(\widehat{\theta}_n^0)\right) = \mathbb{E}_n^{*^{-1}}(xx')\mathbb{E}_n^*(xy) - \mathbb{E}_n^{*^{-1}}(xx')\mathbb{E}_n(xx')\left(\widehat{\theta}_n - \widehat{\theta}_n^0\right) \,,$$

where $\mathbb{E}_n^*$ denotes expectation with respect to the bootstrap distribution P_n^*. The bootstrap QLR statistic is

$$\begin{aligned} QLR_n^* &= 2\,n\left[\widehat{Q}_n^*(\widehat{\theta}_n^{0*}) - \widehat{Q}_n^*(\widehat{\theta}_n^*)\right] = n\,\left(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*}\right)\mathbb{E}_n^*(xx')\left(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*}\right)' \\ &= \left(d'\left(\widehat{\theta}_n^* - h\right)\right)'\left(d'\mathbb{E}_n^{*^{-1}}(xx')d\right)^{-1} d'\left(\widehat{\theta}_n^* - h\right) \,. \end{aligned}$$

It is similar to a Wald statistic, but based on an estimator $\widehat{\theta}_n^*$ that behaves as if H_0 was true.

Quantile Regression. Our assumptions allow for a non-differentiable criterion. For quantile regression of order τ, $q(y, x, \theta) = \rho_\tau(y - x'\theta)$, with $\rho_\tau(u) = u\,(\tau - \mathbb{I}\,(u < 0))$. In practice, estimation is performed solving the dual problem

$$\max_t \left\{y't \middle| - X't = \mathbf{0}, t \in [\tau - 1, \tau]^n\right\} ,$$

where X the matrix of observations on covariates, see Koenker (2005). At the optimum, $t_i(\widehat{\theta}_n) = \tau$ if $y_i - x_i'\widehat{\theta}_n$ is positive, $t_i(\widehat{\theta}_n) = \tau - 1$ if it is negative, and the remaining components are determined so that $-X't(\widehat{\theta}_n) = \mathbf{0}$. Theses value are, up to a constant, the rank scores of the quantile regression, see Gutenbrunner and Jureckova (1992). The dual of the restricted quantile regression writes

$$\max_t \left\{(y - Xh)'t \middle| - H'X't = 0, t \in [\tau - 1, \tau]^n\right\} ,$$

where H is a $p \times (p - r)$ full rank matrix such that $d'H = \mathbf{0}$.[1]

One issue with quantile regression is that the subgradient of the function $\rho_\tau(\cdot)$ can be arbitrarily defined at 0, yielding some indeterminacy. Asymptotically this should be irrelevant, however we have found in practice that using the rank scores to define $S_n(\cdot)$ gives better empirical results. Namely, we used $S_n(\widehat{\theta}_n) = -X't(\widehat{\theta}_n)/n$ and $S_n(\widehat{\theta}_n^0) = -X't(\widehat{\theta}_n^0)/n$. These fulfill the usual properties encountered for a differentiable criterion, namely $S_n(\widehat{\theta}_n) = \mathbf{0}$ and $H'S_n(\widehat{\theta}_n^0) = \mathbf{0}$. That is, the empirical score is zero for components of the parameter space that are unconstrained.

We note that our bootstrap modified criterion can be easily computed following the same lines. Indeed the modified optimization program

$$\min_\theta \sum_{i=1}^n \rho_\tau\left(y_i^* - x_i^{*\prime}\theta\right) - S_n'(\widehat{\theta}_n^0)\theta$$

[1]This comes from rewriting the null hypothesis as $\theta_\dagger = H\gamma_\dagger + h$.

writes in dual form

$$\max_t \left\{ y^{*'} t | - X^{*'} t = n\, S_n(\widehat{\theta}^0_n), t \in [\tau - 1, \tau]^n \right\} .$$

3 QLR Test under Misspecification

3.1 Asymptotic Test

We here state results under misspecification with i.i.d. data without imposing differentiability of the criterion. We consider the following assumption and definition.

Assumption A *(a) Θ is open convex, $\theta_\dagger$ and $\theta^0_\dagger$ as defined in (1.1)–(1.2) are unique. (b) $Q_n(\cdot)$ is convex. (c) (2.3) holds with $A(\theta)$ symmetric positive definite for all θ and $\mathbb{E}\, D(z, \theta_\dagger) = \mathbf{0}$.*

(d) $n^{-1} \sum_{i=1}^n D(z_i, \theta + t) = n^{-1} \sum_{i=1}^n D(z_i, \theta) + A(\theta)t + r_n(t) \quad r_n(t) = o_p(\|t\|) \quad \forall \theta, t.$

(e) $n^{-1/2} \sum_{i=1}^n (D(z_i, \theta) - \mathbb{E}\, D(z, \theta)) \xrightarrow{d} N(0, B(\theta)) \quad \forall \theta \in \Theta.$

(f) d is full rank.

Definition 1 *Let $U = (U_1, \dots U_m)'$ be a vector of m independent standard normal variables and $\lambda = (\lambda_1, \dots \lambda_m) \in R^m$. The random variable $\sum_{i=1}^m \lambda_m U_m^2$ is distributed as a weighted sum of chi-squares with parameters (m, λ) and its distribution is denoted by $M_m(\cdot, \lambda)$.*

Theorem 3.1 *For a random sample $\{z_i, i = 1, \dots n\}$, under Assumption A,*

(a) Under H_0, $\mathrm{QLR}_n \xrightarrow{d} M_r(\cdot, \lambda)$, where λ is the vector of r eigenvalues of

$$\left(d' A_\dagger^{-1} d\right)^{-1} d' A_\dagger^{-1} B_\dagger A_\dagger^{-1} d \,, \qquad \text{where } A_\dagger = A\left(\theta_\dagger\right) \text{ and } B_\dagger = B\left(\theta_\dagger\right) .$$

(b) Under H_A, $\mathrm{QLR}_n / n \xrightarrow{p} c > 0$.

When the second Bartlett's identity holds, that is $B_\dagger = A_\dagger$, λ is a vector of ones and the QLR statistic asymptotically follows a centered chi-square distribution with r degrees of freedom. In a correctly specified linear regression, $B_\dagger = \sigma^2 A_\dagger$ for homoscedastic errors with variance σ^2, and we can easily render QLR asymptotically pivotal. In general however, the QLR statistic has a more involved asymptotic distribution under H_0. Our general characterization of λ appears to be new, see Lien and Vuong (1987, Lemma 2) for an alternative one in the linear regression model.

3.2 Robust QLR Test

We here focus on an univariate restriction for which one can generally build a robust QLR statistic since λ is scalar. This is of particular empirical relevance for testing the significance of a specific parameter component. Given a consistent estimator $\widehat{\lambda}_n$, the robust QLR statistic is

$$\mathrm{RQLR}_n = \frac{\mathrm{QLR}_n}{\widehat{\lambda}_n} \xrightarrow{d} \chi^2_1 \qquad \text{under } H_0\,,$$

and a test can be entertained using standard critical values. Estimating λ requires consistent estimation of $A_\dagger^{-1}$ and $B_\dagger$. One can use empirical analogs of the above matrices evaluated at a consistent estimator of $\theta_\dagger$. The quantity $A_\dagger^{-1} B_\dagger A_\dagger^{-1}$ happens to be the asymptotic variance of $\widehat{\theta}_n$, which should be estimated to build the robust Wald and Score statistics. Hence it is as easy (or as difficult) to build the robust QLR statistic than it is to get robust Wald or Score statistics.

For testing a single restriction $d'(\theta_\dagger - h) = 0$ in linear regression, an estimator of λ is easily obtained as

$$\widehat{\lambda}_n = \frac{d'\mathbb{E}_n^{-1}(xx')\Sigma_n \mathbb{E}_n^{-1}(xx')d}{d'\mathbb{E}_n^{-1}(xx')d} \quad \text{with} \quad \Sigma_n = \mathbb{E}_n(xx'(y - x'\widehat{\theta}_n)^2) = n^{-1}\sum_{i=1}^n x_i x_i'(y_i - x_i'\widehat{\theta}_n)^2\,,$$

see Eicker (1967) and White (1982). The numerator is the heteroscedasticity-robust variance estimator provided by most statistical software.

For quantile regression, $A_\dagger = \mathbb{E}\left(f_\varepsilon(0|x)xx'\right)$ depends on the conditional density of the error term $\varepsilon = y - x'\theta_\dagger$. Powell (1991) proposes to use the estimator

$$\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{y_i - x_i'\widehat{\theta}_n}{h}\right) x_i x_i' \tag{3.5}$$

where $K(\cdot)$ is a density and h a bandwidth. Also $B_\dagger = \mathbb{E}\left(\left(\tau - \mathbb{I}(\varepsilon \leq 0)\right)^2 xx'\right)$ can be estimated by the Eicker-type estimator

$$\frac{1}{n}\sum_{i=1}^{n}\left(\tau - \mathbb{I}(y_i - x_i'\widehat{\theta}_n \leq 0)\right)^2 x_i x_i',$$

see Kim and White (2003).

3.3 Bootstrap Test

Our bootstrap test rejects H_0 if $\text{QLR}_n > q^*_{1-\alpha}$, where $q^*_{1-\alpha}$ is the $1-\alpha$ quantile of QLR^*_n. We first establish consistency of our bootstrap procedure.

Theorem 3.2 *Under Assumption A,* QLR^*_n *is bounded in probability conditionally on the sample, and under* H_0, $\Pr\left(\text{QLR}^*_n \leq x \,\middle|\, (y_i, x_i)_{i=1,\ldots,n}\right) - \Pr\left(\text{QLR}_n \leq x\right) = o_p(1)$.

As QLR^*_n is bounded in probability, it would be straightforward to show that our bootstrap test has non trivial power under $\sqrt{n}$-local alternatives.

For an univariate restriction, we can obtain an asymptotically pivotal statistic, so we can hope to obtain asymptotic refinements. The next result confirms this holds for linear regression, irrespective of whether the model is well specified.[2] The only competing

[2] We focus on least-squares regression, as the Edgeworth expansion for quantiles is typically non-standard due to the lattice nature of the subgradient, see e.g. Falk and Janas (1992).

method that has been shown to yield asymptotic refinements under misspecification is the nonparametric bootstrap Wald test, see Hall and Horowitz (1996). From our previous results,

$$\mathrm{RQLR}_n = \frac{\mathrm{QLR}_n}{\widehat{\lambda}_n} = n\left(\widehat{\theta}_n - h\right)' d\left(d'\mathbb{E}_n^{-1}(xx')\Sigma_n\mathbb{E}_n^{-1}(xx')d\right)^{-1} d'\left(\widehat{\theta}_n - h\right),$$

and similarly

$$\mathrm{RQLR}_n^* = \frac{\mathrm{QLR}_n^*}{\widehat{\lambda}_n^*} = \left(\widehat{\theta}_n^* - h\right)' d\left(d'\mathbb{E}_n^{*-1}(xx')\Sigma_n^*\mathbb{E}_n^{*-1}(xx)d\right)^{-1} d'\left(\widehat{\theta}_n^* - h\right),$$

where $\Sigma_n^* = \mathbb{E}_n^*(xx'(y - x'\widehat{\theta}_n^*)^2) = n^{-1}\sum_{i=1}^n x_i^* x_i^{*\prime}(y_i^* - x_i^{*\prime}\widehat{\theta}_n^*)^2$.

Theorem 3.3 *Under Assumption A and*

(a) the Cramer condition $\overline{\lim}_{t\to\infty}\left|\mathbb{E}\exp\left(i\,t\,d'\mathbb{E}^{-1}(xx')x\,(y - x'\theta_\dagger)\right)\right| < 1$*,*

(b) $\mathbb{E}\left(d'\mathbb{E}^{-1}(xx')xy\right)^{12} < \infty$ *and* $\mathbb{E}\left(d'\mathbb{E}^{-1}(xx')xx'\theta_\dagger\right)^{12} < \infty$,

$$\Pr\left(\mathrm{RQLR}_n^* \le x \,\middle|\, (y_i, x_i)_{i=1,\dots,n}\right) - \Pr\left(\mathrm{RQLR}_n \le x\right) = O_p(n^{-3/2}) \qquad \text{under } H_0\,.$$

Moreover if $q_{R,1-\alpha}^*$ *is the* $1-\alpha$ *quantile of the bootstrap distribution of* RQLR_n^**,*

$$P\left(\mathrm{RQLR}_n \le q_{R,1-\alpha}^*\right) = 1 - \alpha + O(n^{-2}) \qquad \text{under } H_0\,.$$

The Cramer condition automatically holds if some components of x have an absolutely continuous part. It is minimal to ensure that the influence function of the functional appearing in the quadratic form of QLR_n is non lattice. The condition could be replaced by the multivariate Cramer condition $\overline{\lim}_{||u||\to\infty}|\mathbb{E}\exp(iu'x(y - x'\theta_\dagger)| < 1$. Moment conditions of order 12 ensure that the terms in the Edgeworth expansion of RQLR_n depending on moments of order up to 6 match the empirical ones in the bootstrap distribution up to $O_p(n^{-1/2})$ by a standard CLT argument. Such conditions are implied by the existence of higher moments for both y and x.

4 Small Sample Evidence and Application

4.1 Simulations

We focused on inference on the coefficient β_2 in the assumed linear model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \,.$$

We generated the variable y according to the model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \psi x_{1i} x_{2i} + (1 + l|x_{2i}|)\, \eta_i \,,$$

where $(\beta_1, \beta_2) = (0, 0)$. The variable x_1 is standard Gaussian, x_2 is independent lognormal with mean 0 and variance 1, and η is independently distributed as a Student law. This specification provides us with an asymmetric covariate and generates observations with high leverage, which can create serious obstacles to heteroscedasticity robust inference as shown by Chesher and Austin (1991).

The parameter l controls heteroscedasticity. When $\psi \neq 0$, the linear conditional mean (or quantile) is misspecified. Due to the regressors' independence, β_2 is unchanged under misspecification of the conditional mean or median, which is convenient when studying the tests' behavior for $\psi \neq 0$. We considered values of $l = 0.5$ and $\psi = 0.5$, which corresponds to moderate misspecifications that could go unnoticed.

We ran 20000 simulations for $n = 200$. To speed up computations, we use the warp-speed method proposed by Davidson and MacKinnon (2007) and studied by Giacomini, Politis, and White (2013). Specifically, for each considered null hypothesis, we drew one bootstrap and double bootstrap sample for each simulated data, and we used the whole set of bootstrap statistics to compute the bootstrap and double bootstrap p-values associated with each original statistic. We report our results using graphs. The first type of graph draws the errors in rejection probability (ERP), that is the difference between nominal size and the empirical rejection proportion under the null hypothesis. A perfect test would exhibit an ERP of zero for any nominal size. This gives us a visual way to evaluate whether the null distribution of the test statistic is well approximated by its asymptotic or bootstrap approximation. The second type of graph draws the power curves of each test.

Mean Regression We set $\eta \sim t_5$. The QLR statistic was bootstrapped under the null (denoted by QLR0-b) and we used double bootstrap to obtain improvements (denoted by QLR0-db). We compare them to the usual bootstrap with naive resampling and testing $\beta_2 = \widehat{\beta}_2$ in the bootstrap world (QLR-b and QLR-db). We also consider RQLR, the robust QLR statistic equal to the Wald one. We applied bootstrapping under the null (RQLR0-b) as well as the usual nonparametric bootstrap (RQLR-b), see Hall and Horowitz (1996). When using robust statistics, we estimated the correction factor λ by the HC3 method, as recommended by Long and Ervin (2000) and Cribari-Neto, Ferrari, and Oliveira (2005).

Figure 1 gathers our results for size control. The robust asymptotic test is always undersized at usual nominal sizes, and bootstrap very imperfectly corrects this phenomenon. The bootstrap test QLR-b is always oversized at usual nominal sizes and double bootstrap is moderately helpful. Double bootstrap under the null always yields the best size control.

Figure 2 gathers our results in terms of power. Tests that are systematically oversized have a slightly better power, as could be expected, while other tests have similar performances.

Quantile Regression Since most existing bootstrap methods for inference in quantile regression assume a correct model, see Kocherginsky, He, and Mu (2005) for a review, we

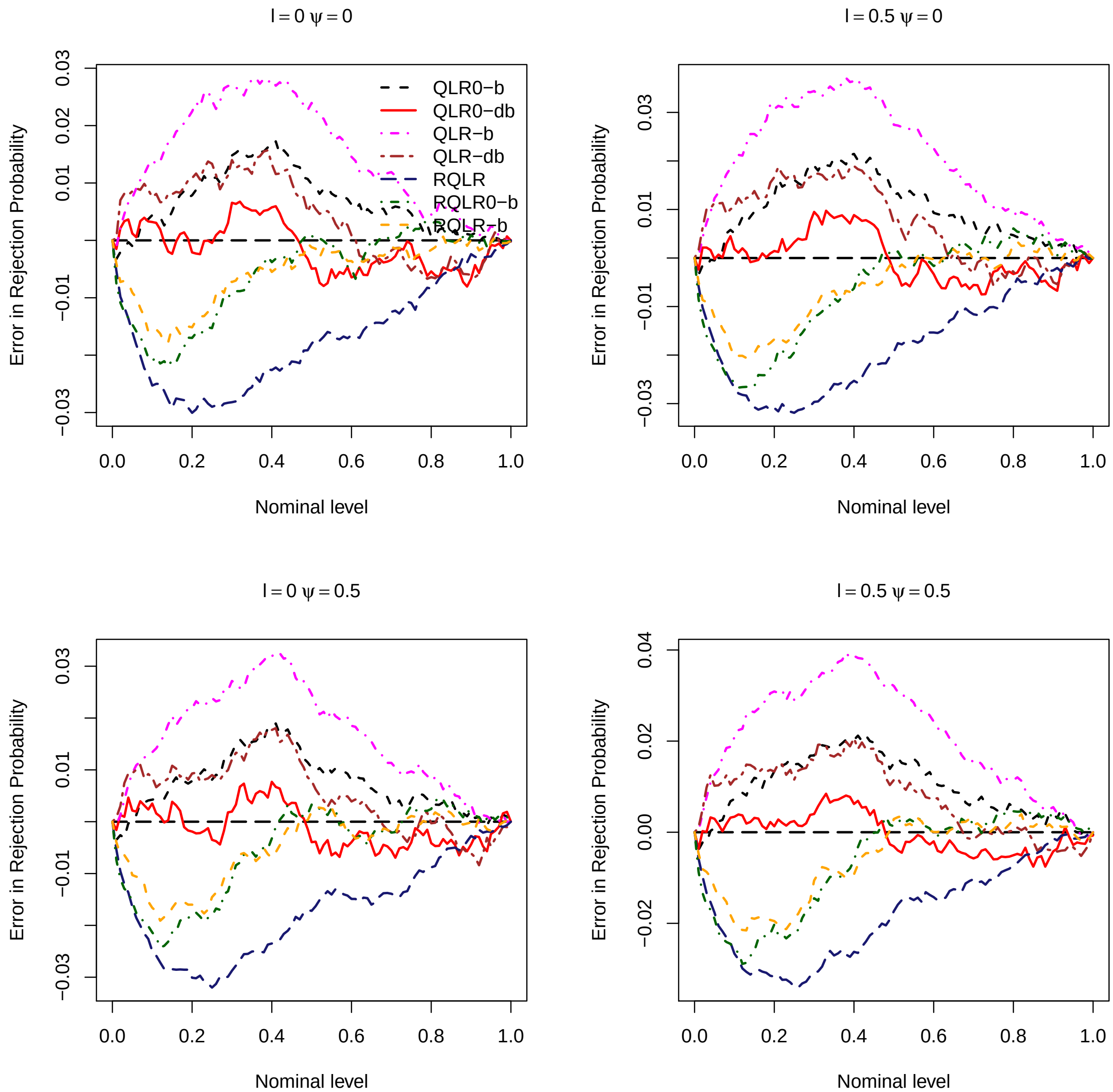

Figure 1: Errors in Rejection Probabilities for mean regression with $\eta \sim t_5$.

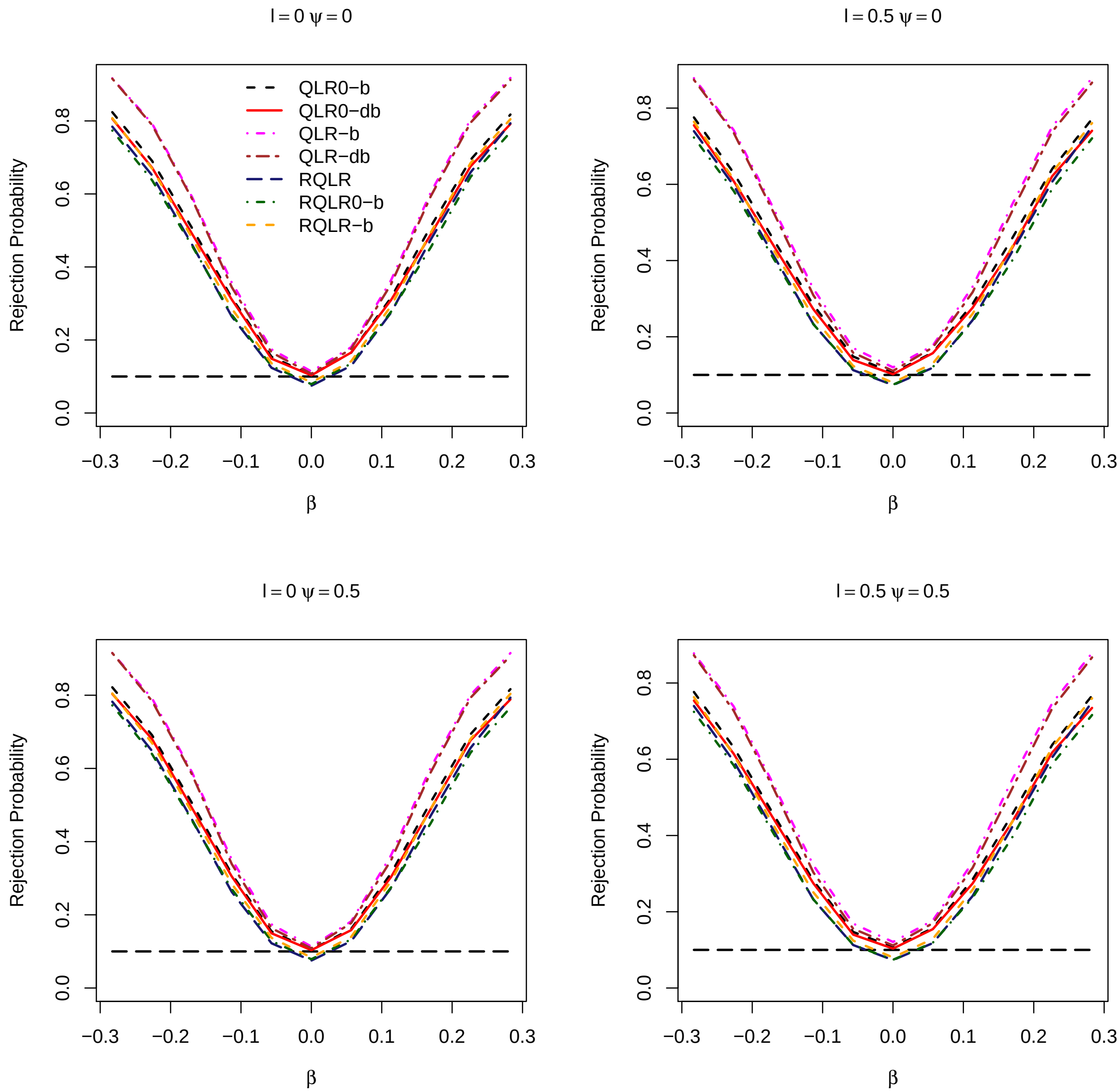

Figure 2: Power curves for mean regression with $\eta \sim t_5$.

only compared five tests. The first two are based on robust asymptotic and bootstrap Wald tests, denoted as W and W-b. For standard errors, we used the formula detailed in Section 3.2, and specifically we chose $K(\cdot)$ as the standard normal density and $h = 0.79\, n^{-1/5}\, \mathrm{IQR}$ in (3.5). We also consider the percentile bootstrap, denoted as P-b, which was found to be the best performing method by Tarr (2012). We compare these to our bootstrap under the null QLR test, denoted as QLR0-b, and its double bootstrap version, QLR0-db. We did not look at bootstrap with naive resampling for testing $\beta_2 = \widehat{\beta}_2$ in the bootstrap world, given its poor performances in our previous experiments. We considered several setups corresponding to median regression and quantile regression of order $\tau = 0.25$, and η distributed as t_5 or t_1. We report a selection of our results.

Figures 3 and 5 gather our results for size control in median regression, while Figure 7 consider quantile regression of order 0.25. Under misspecification, the asymptotic robust Wald test does not perform well, and bootstrap is not very successful at correcting this phenomenon. The percentile bootstrap performs better, but double bootstrap under the null of the QLR test provides the best size control.

Figures 4, 6, and 7 reports power curves. Misspecification, and in particular heteroscedasticity, has very adverse effects on the tests' performances. Under misspecification, Wald tests can have an erratic behavior, leading to a sometimes non-monotone power. Besides the Wald test, QLR-0db has always highest power under misspecification.

4.2 Empirical Application.

We considered some parametric quantile models for children birthweight using data analyzed by Abrevaya (2001) and Koenker (2005), who also gave a detailed data description. We focused on median regression and lower quantiles regression on a subsample of 1089

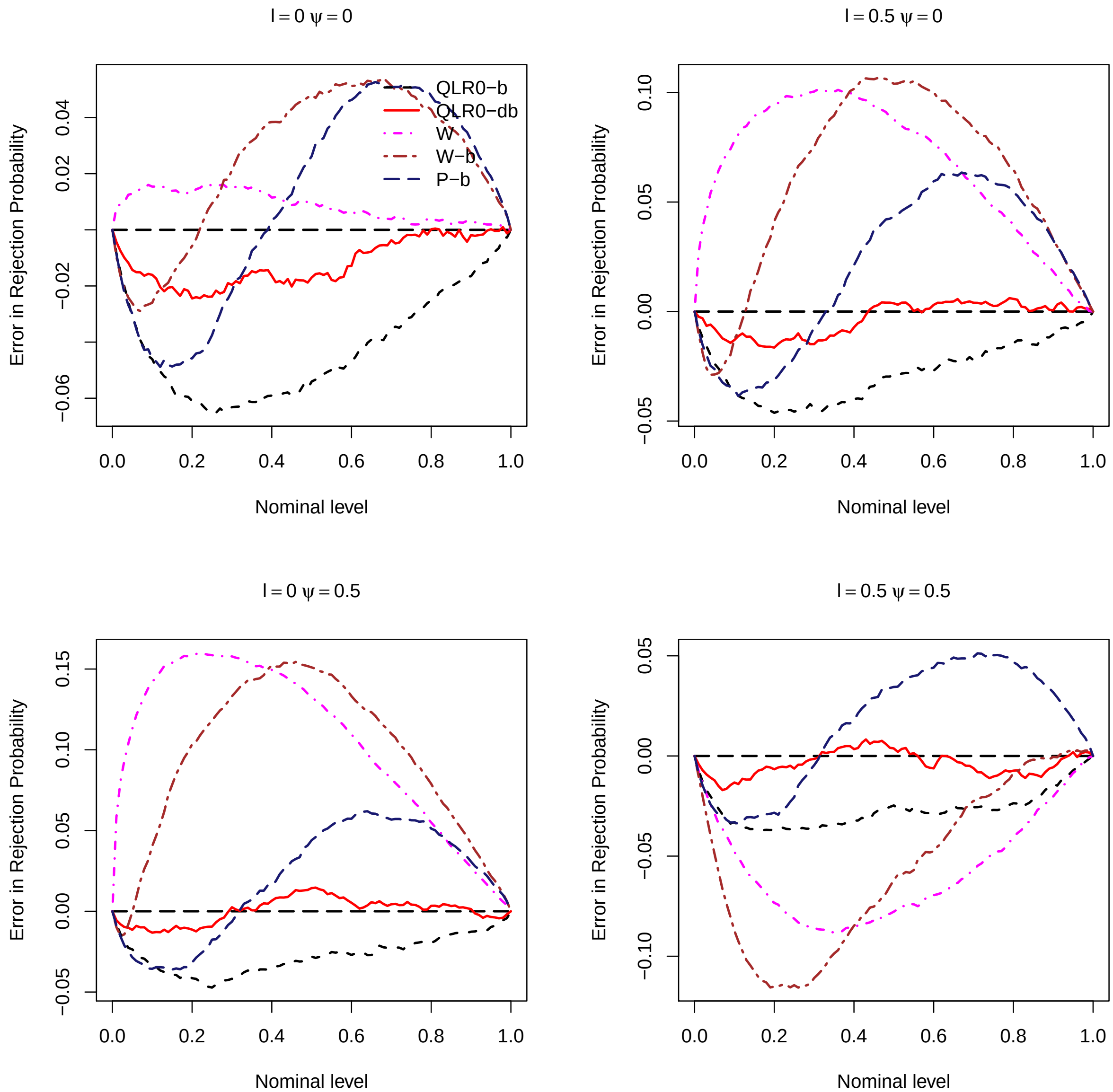

Figure 3: Errors in rejection probabilities for median regression with $\eta \sim t_5$.

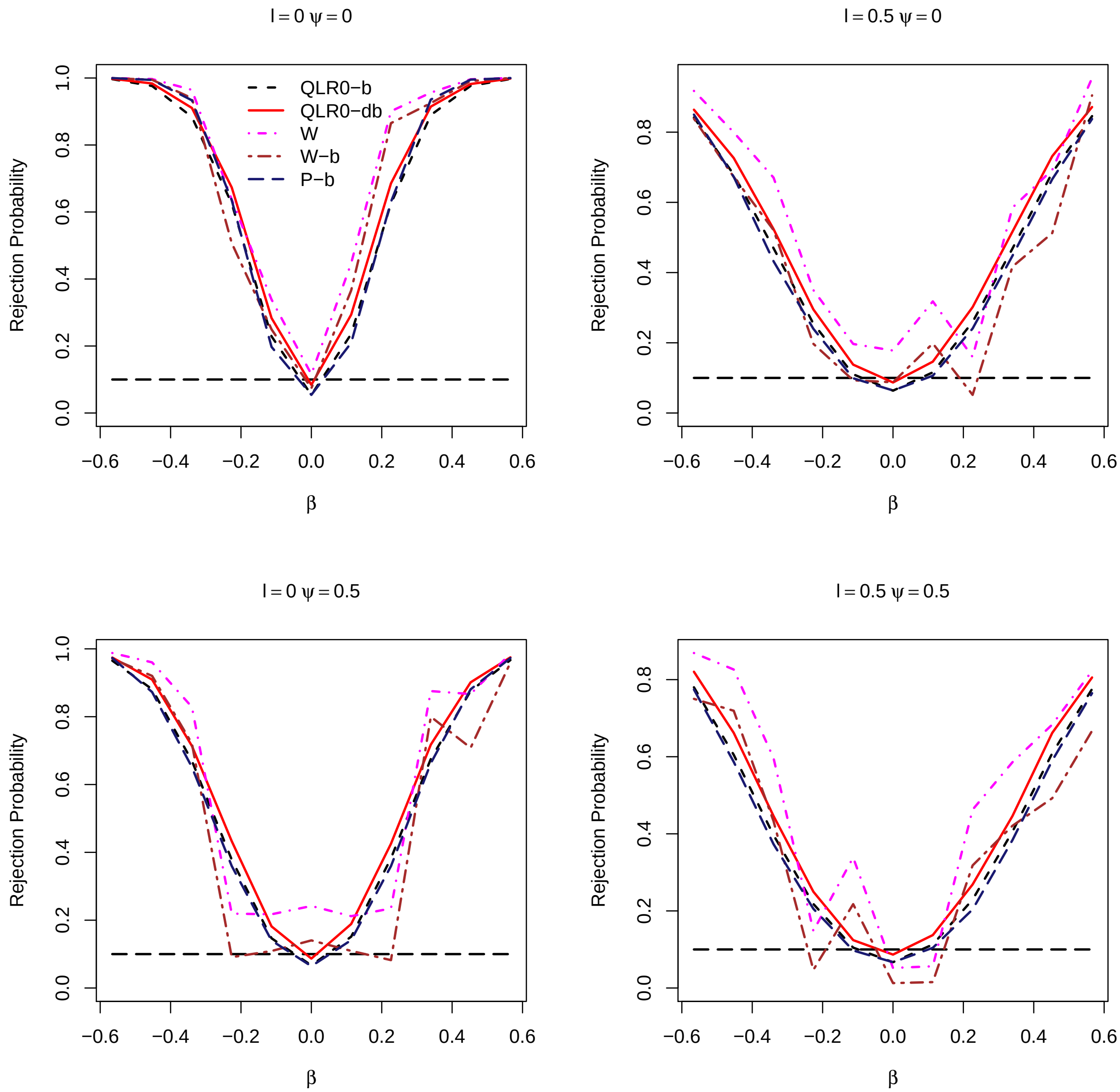

Figure 4: Power curves for median regression with $\eta \sim t_5$.

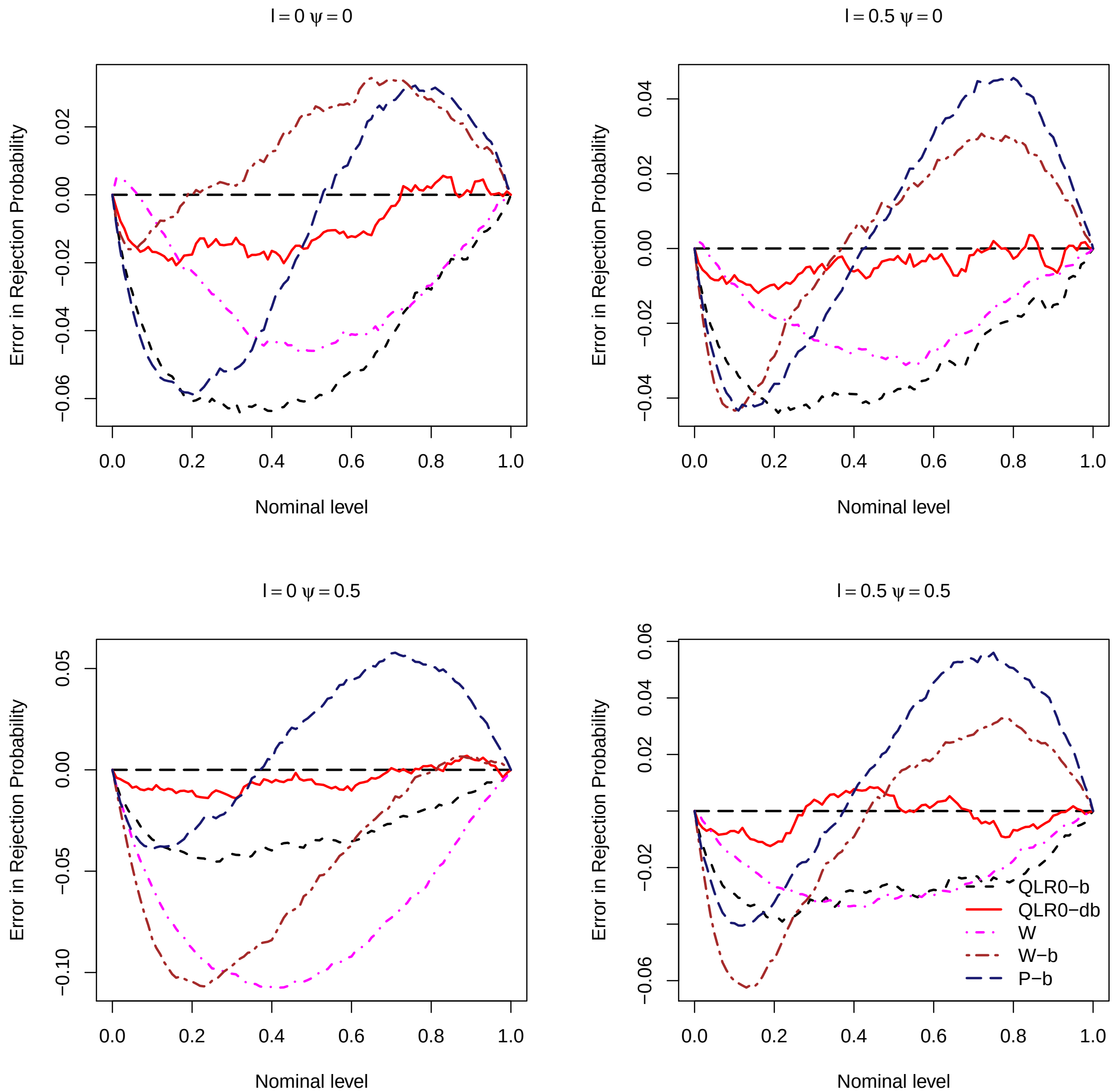

Figure 5: Errors in rejection probabilities for median regression with $\eta \sim t_1$.

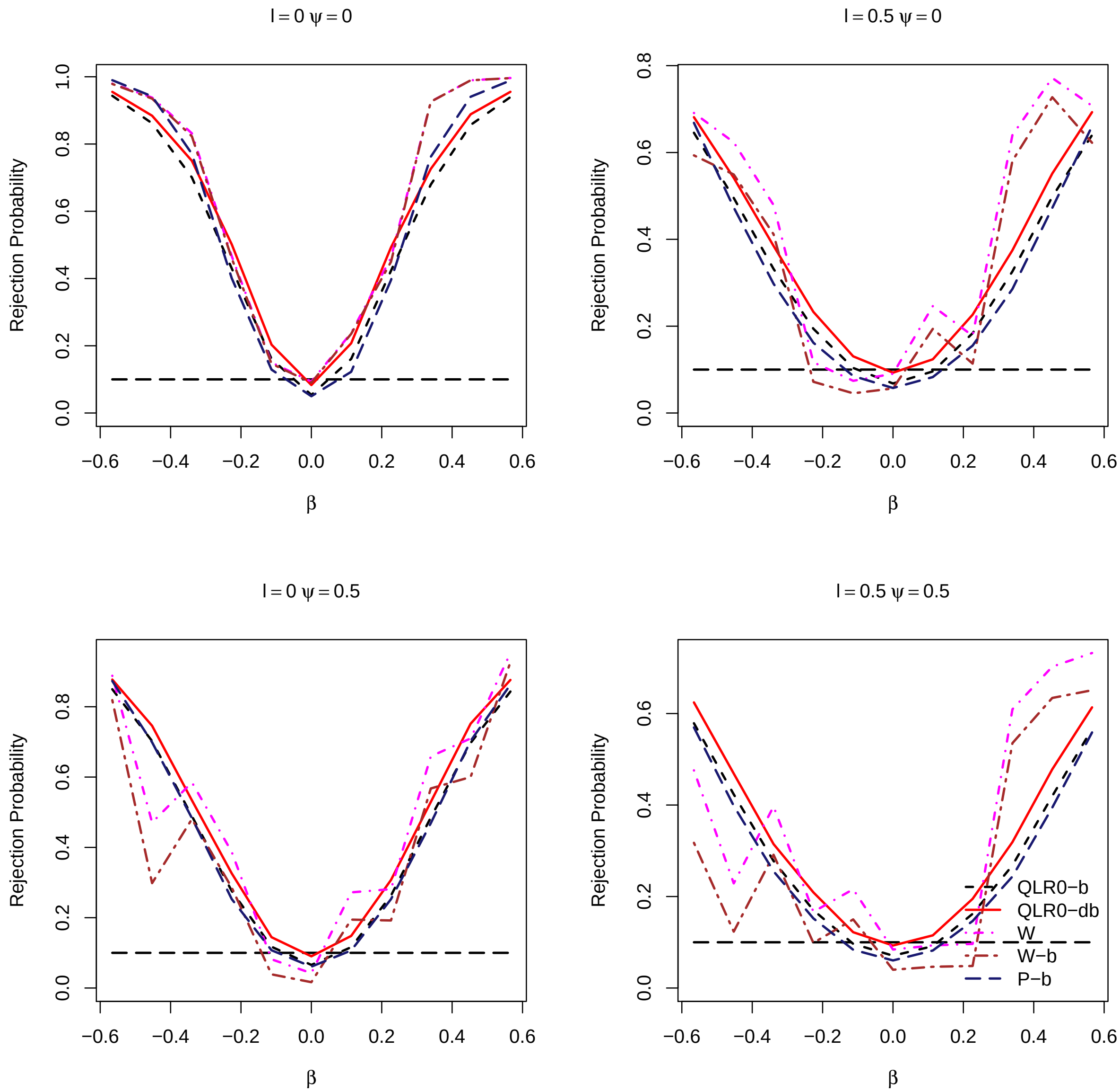

Figure 6: Power curves for median regression with $\eta \sim t_1$.

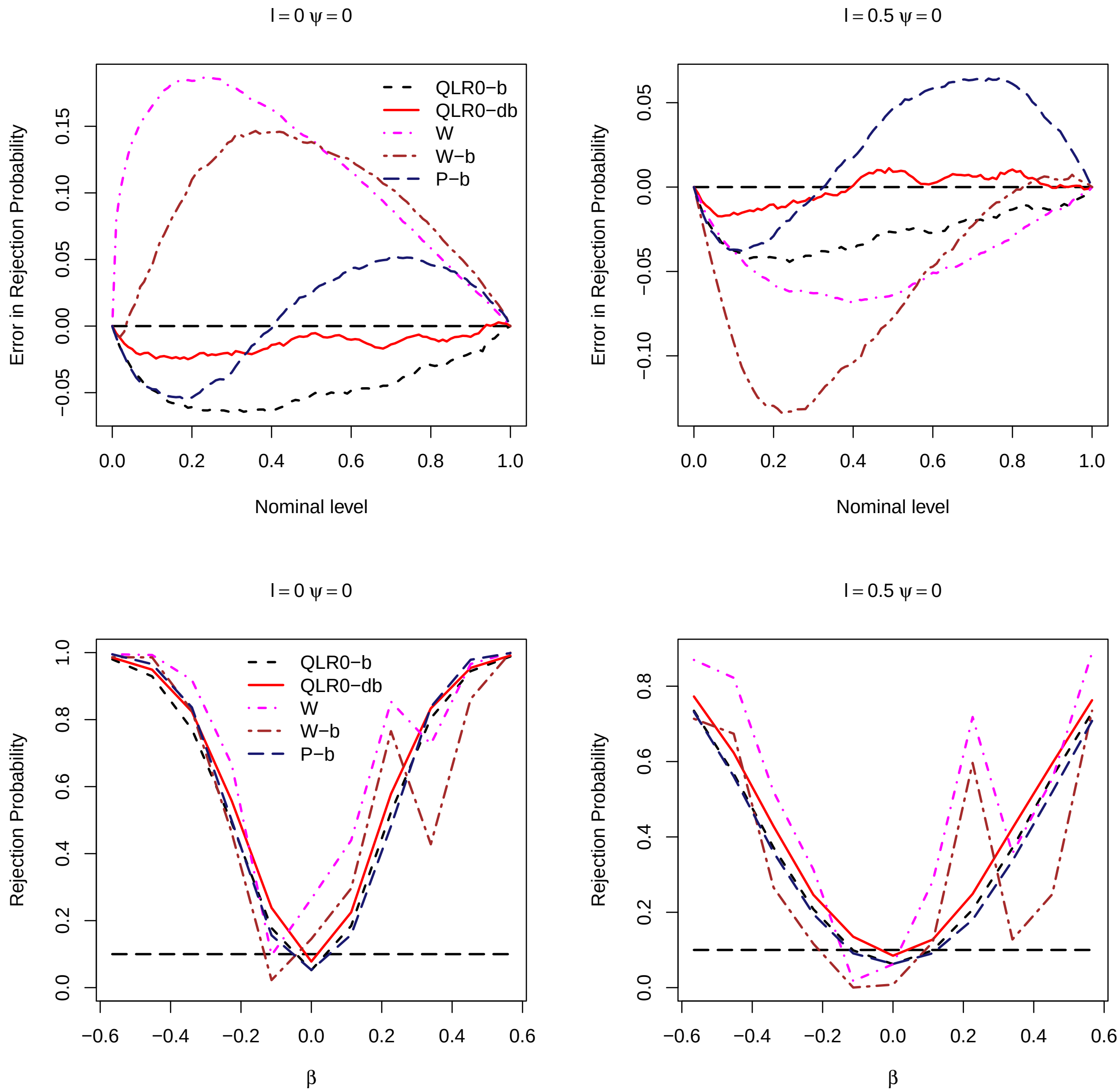

Figure 7: Errors in rejection probabilities and power curves for quantile regression with $\tau = .25$ and $\eta \sim t_5$.

Table 1: 90% Confidence Intervals for CIGAR parameter

	$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$
Wald	[-13.83 , 1.49]	[-9.65 , -0.54]	[-9.46 , -1.31]
Wald-b	[-17.09 , 4.67]	[-10.14 , -0.07]	[-9.93 , -0.86]
P-b	[-15.60 , 3.15]	[-10.00 , -0.03]	[-9.83 , -0.94]
QLR0-b	[-15.72 , 1.73]	[-9.86 , 0.23]	[- 8.95 , -0.93]
QLR0-db	[-16.06 , 1.73]	[-9.87 , 0.41]	[-8.95 , -0.74]

smoking college graduate white mothers. We considered a model linear in (i) weight gain during pregnancy, (ii) average number of cigarettes per day (denoted as CIGAR), and (iii) dummies indicating whether the child is male and whether the mother is married, and quadratic in mother's age as suggested by Koenker (2005). Maistre, Lavergne, and Patilea (2017) reported that there may be misspecification for lower quantiles, but did not find evidence of misspecification in median regression. In Table 1, we report 90% confidence intervals for the CIGAR parameter using the same methods as in our simulations. We considered 999 bootstrap samples for each potential value of the parameter.

For median regression, there are little differences between the outcomes of different methods, but confidence intervals from QLR0-b and QLRO-db are among the shortest, while the bootstrap Wald and percentile intervals are largest. For lower quantiles, the asymptotic Wald-based confidence interval is mostly shortest, while the bootstrap one is much larger. This is coherent with our simulation findings: Wald test does not control size well and can be severely oversized, and the bootstrap makes it more conservative. The difference between confidence intervals can be so large that for $\tau = 0.25$ the Wald and percentile bootstrap signals a significant effect of cigarettes consumption on birthweight, while the confidence intervals based on QLR indicates a non-significant effect. For $\tau = 0.1$, our bootstrap method delivers tighter intervals than competing bootstrap methods, which is likely related to its superior power performances under misspecification as observed in simulations.

5 Conclusion

We propose a simple bootstrap method, bootstrapping under the null, for quasi likelihood-ratio tests that provides valid critical values even under misspecification. A key advantage of QLR tests is that they do not necessitate to estimate a robust-covariance matrix: this can be difficult, as in quantile regression, and can severely affect size control and power performances. We found that our method yield in practice rejection probabilities that are close to nominal levels in small samples, and double bootstrapping under the null the non-robust QLR is preferable to relying on a robust version for size control as well as power. We considered an i.i.d. setting and a particular class of convex estimation criteria, but it is likely that the same idea could be extended to dependent data and more general criteria. We intend to investigate along these lines in future work.

6 Proofs

We first recall two useful results.

Theorem 6.1 (Andersen and Gill (1982, Theorem II.1)) *Let $Q_n(\cdot)$ be a sequence of random convex functions defined on an open convex Θ such that $Q_n(\theta) \xrightarrow{p} Q(\theta)$ for any*

$\theta \in \Theta$. *Then for any compact subset K of Θ,* $\sup_{\theta \in K} |Q_n(\theta) - Q(\theta)| \xrightarrow{p} 0$.

Theorem 6.2 (Hjort and Pollard (1993, Basic Corollary)) *Suppose $A_n(s)$ is convex and can be represented as $\frac{1}{2}s'Vs + U_n's + C_n + r_n(s)$, where V is symmetric and positive definite, U_n is stochastically bounded, C_n is arbitrary, and $r_n(s) = o_p(1)$ for each s. Then α_n, the argmin of $A_n(s)$, is only $o_p(1)$ away from $\beta_n = -V^{-1}U_n$, the argmin of $\frac{1}{2}s'Vs+U_n's$. If also $U_n \xrightarrow{d} U$ then $\alpha_n \xrightarrow{d} -V^{-1}U$.*

Proof of Theorem 3.1. (a) Let $\Delta_n = n^{-1/2}\sum_{i=1}^n D_i(z_i, \theta_\dagger)$. Write

$$n\left[Q_n(\theta_\dagger + t/\sqrt{n}) - Q_n(\theta_\dagger)\right] = \Delta_n' t + \frac{1}{2}t'At + nR_n(t/\sqrt{n}),$$

with $nR_n(t/\sqrt{n}) = o_p(\|t\|^2)$ for all t. By Theorem 6.2, $\sqrt{n}(\widehat{\theta}_n - \theta_\dagger) = -A_\dagger^{-1}\Delta_n + o_p(1)$.

Rewrite the null hypothesis as $H_0: \ \theta_\dagger = H\gamma_\dagger + h$, then under H_0

$$\begin{aligned} n\left[Q_n(H(\gamma_\dagger + u/\sqrt{n}) + h) - Q_n(H\gamma_\dagger + h)\right] &= \Delta_n' Hu + \frac{1}{2n}u'H'A_\dagger Hu + nR_n(Hu/\sqrt{n}) \\ &\Rightarrow \sqrt{n}(\widehat{\theta}_n^0 - \theta_\dagger) = -H(H'A_\dagger H)^{-1}H'\Delta_n + o_p(1). \end{aligned}$$

From the quadratic approximation of $Q_n(\theta_\dagger)$,

$$\begin{aligned} \mathrm{QLR}_n &= n(\widehat{\theta}_n - \theta_*)'A_\dagger(\widehat{\theta}_n - \theta_*) - n(\widehat{\theta}_n^0 - \theta_\dagger)'A_\dagger(\widehat{\theta}_n^0 - \theta_\dagger) + o_p(1) \\ &= n(\widehat{\theta}_n - \widehat{\theta}_n^0)'A_\dagger(\widehat{\theta}_n - \widehat{\theta}_n^0) + o_p(1), \end{aligned}$$

because $(\widehat{\theta}_n - \widehat{\theta}_n^0)'A_\dagger(\widehat{\theta}_n^0 - \theta_\dagger) = o_p(1)$. This is a quadratic form in the vector

$$\sqrt{n}(\widehat{\theta}_n - \widehat{\theta}_n^0) = \left[A_\dagger^{-1} - H(H'A_\dagger H)^{-1}H'\right]\Delta_n + o_p(1).$$

It thus converges to a weighted sum of independent chi-squares with one degree of freedom. From Vuong (1989, Lemma 3.2), the vector λ contains the eigenvalues of

$$B_\dagger\left[A_\dagger^{-1} - H(H'A_\dagger H)^{-1}H'\right]A_\dagger\left[A_\dagger^{-1} - H(H'A_\dagger H)^{-1}H'\right] = B_\dagger A_\dagger^{-1/2}MA_\dagger^{-1/2}.$$

As M is idempotent, the above has the same eigenvalues as $MA_\dagger^{-1/2}B_\dagger A_\dagger^{-1/2}M$. Now

$$C = \begin{bmatrix} (H'A_\dagger H)^{-1/2} H'A_\dagger^{1/2} & \mathbf{0} \\ \mathbf{0} & \left(d'A_\dagger^{-1}d\right)^{-1/2} d'A_\dagger^{-1/2} \end{bmatrix}$$

is an orthogonal matrix. Hence W has the same eigenvalues as $CMA_\dagger^{-1/2}B_\dagger A_\dagger^{-1/2}MC'$, whose non-zero eigenvalues are the ones of $\left(d'A_\dagger^{-1}d\right)^{-1/2} d'A_\dagger^{-1}B_\dagger A_\dagger^{-1}d\left(d'A_\dagger^{-1}d\right)^{-1/2}$. These are unchanged by pre-multiplying by $\left(d'A_\dagger^{-1}d\right)^{-1/2}$ and post-multiplying by its inverse.

(b) Under H_A, $Q_n(\widehat{\theta}_n^0) - Q_n(\widehat{\theta}_n)$ converges to a positive real by Theorem 6.1. ■

Proof of Theorem 3.2. Let $\Delta_n^{0*} = n^{-1/2}\sum_{i=1}^n D_i(z_i^*, \theta_\dagger^0)$. We have

$$n\left[\widehat{Q}_n^*(\theta_\dagger^0 + t/\sqrt{n}) - \widehat{Q}_n^*(\theta_\dagger^0)\right] = \left(\Delta_n^* - \sqrt{n}S_n(\widehat{\theta}_n^0)\right)' t + \frac{1}{2}t'A_\dagger^0 t + nR_n(t/\sqrt{n}),$$

where $nR_n(t/\sqrt{n}) = o_p(\|t\|^2)$ for each t and $A_\dagger^0 = A\left(\theta_\dagger^0\right)$. From Theorem 6.2, the minimizer of $\widehat{Q}_n^*(\theta_\dagger^0 + t/\sqrt{n})$ is

$$\sqrt{n}(\widehat{\theta}_n^* - \theta_\dagger^0) = -\left(A_\dagger^0\right)^{-1}\left(\Delta_n^* - \sqrt{n}S_n(\widehat{\theta}_n^0)\right) + o_p(1).$$

A similar reasoning shows that

$$\begin{aligned}
\sqrt{n}(\widehat{\theta}_n^{0*} - \theta_\dagger^0) &= -H(H'A_\dagger^0 H)^{-1}H\left(\Delta_n^* - \sqrt{n}S_n(\widehat{\theta}_n^0)\right) + o_p(1) \\
\text{QLR}_n^* &= n(\widehat{\theta}_n^* - \theta_\dagger^0)'A_\dagger^0(\widehat{\theta}_n^* - \theta_\dagger^0) - n(\widehat{\theta}_n^{0*} - \theta_\dagger^0)'A_\dagger^0(\widehat{\theta}_n^{0*} - \theta_\dagger^0) + o_p(1) \\
&= n(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*})'A_\dagger^0(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*}) + o_p(1),
\end{aligned}$$

because $(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*})'A_\dagger^0(\widehat{\theta}_n^{0*} - \theta_\dagger^0) = o_p(1)$. Now

$$\sqrt{n}(\widehat{\theta}_n^* - \widehat{\theta}_n^{0*}) = \left[\left(A_\dagger^0\right)^{-1} - H(H'A_\dagger^0 H)^{-1}H'\right]\left(\Delta_n^* - \Delta_n^0 + \Delta_n^0 - \sqrt{n}S_n(\widehat{\theta}_n^0)\right) + o_p(1),$$

where $\Delta_n^0 = n^{-1/2}\sum_{i=1}^n D_i(z_i, \theta_\dagger^0)$. As $\Delta_n^* - \Delta_n^0$ has conditionally on the initial sample the same asymptotic distribution than $\Delta_n^0 - n^{1/2}\mathbb{E}\, D(z, \theta_\dagger^0)$ by Gine and Zinn (1990), it is bounded in probability. Moreover,

$$\Delta_n^0 - \sqrt{n} S_n(\widehat{\theta}_n^0) = A_\dagger^0 \left(\widehat{\theta}_n^0 - \theta_\dagger^0\right) + o_p\left(\|\widehat{\theta}_n^0 - \theta_\dagger^0\|\right) = O_p(1)\,.$$

Hence QLR_n^* is bounded in probability.

Under H_0, $\theta_\dagger^0 = \theta_\dagger$, $-\left[(A_\dagger)^{-1} - H(H'A_\dagger H)^{-1}H'\right] A_\dagger \sqrt{n}(\widehat{\theta}_n^0 - \theta_\dagger) = o_p(1)$, and QLR_n^* is conditionally on the initial sample asymptotically distributed as $M_r(\cdot, \lambda)$. ■

Proof of Theorem 3.3. We adopt a functional view of the regression parameters, which allows to compute the influence functions of the parameters and prove the asymptotic validity of the method when standardizing by the correct variance. We then derive Edgeworth expansions of the bootstrap and original statistic and show that their coefficients coincide up to terms of order $O_p(n^{-1/2})$. Define

$$\gamma(P) \equiv \gamma = \left(d'\mathbb{E}^{-1}(xx')d\right)^{-1} d'\left(\theta_\dagger - h\right) \quad \text{and} \quad \Sigma = \mathbb{E}\left(xx'(y - x'\theta_\dagger)^2\right)\,.$$

Let $\gamma(P_n) \equiv \gamma_n$ and $\gamma(P_n^*) \equiv \gamma_n^*$ be similar quantities defined by replacing $\mathbb{E}$ by $\mathbb{E}_n$ or $\mathbb{E}_n^*$. Then

$$\mathrm{RQLR}_n = \frac{n\,\gamma_n^2\left(d'\mathbb{E}_n^{-1}(xx')d\right)^2}{d'\mathbb{E}_n^{-1}(xx')\Sigma_n\mathbb{E}_n^{-1}(xx')d}$$
$$\mathrm{RQLR}_n^* = \frac{n\,\left(\gamma_n^* - \gamma_n\right)^2\left(d'\mathbb{E}_n^{*-1}(xx')d\right)^2}{d'\mathbb{E}_n^{*-1}(xx')\Sigma_n^*\mathbb{E}_n^{*-1}(xx')d}\,.$$

Since $\gamma(P)$ is a smooth functional of moments, its influence function is given by

$$\gamma^{(1)}(y,x) = \left(d'\mathbb{E}^{-1}(xx')d\right)^{-1} d'\mathbb{E}^{-1}(xx')x(y - x'\theta_\dagger)$$
$$- d'\left(\theta_\dagger - h\right)\frac{d'\mathbb{E}^{-1}(xx')\left(xx' - \mathbb{E}\,xx'\right)\mathbb{E}^{-1}(xx')d}{\left(d'\mathbb{E}^{-1}(xx')d\right)^2}\,.$$

Under H_0, the asymptotic variance of γ_n is

$$\mathrm{Var}\left(\gamma^{(1)}(y,x)\right) = \frac{d'\mathbb{E}^{-1}(xx')\,\Sigma\,\mathbb{E}^{-1}(xx')d}{\left(d'\mathbb{E}^{-1}(xx')d\right)^2} = \omega^2 .$$

Define the empirical counterpart and its bootstrap version as

$$\omega_n^2 = \frac{d'\mathbb{E}_n^{-1}(xx')\,\Sigma_n\,\mathbb{E}_n^{-1}(xx')d}{\left(d'\mathbb{E}_n^{-1}(xx')d\right)^2} \quad \text{and} \quad \omega_n^{2*} = \frac{d'\mathbb{E}_n^{*-1}(xx')\Sigma_n^*\,\mathbb{E}_n^{*-1}(xx')d}{\left(d'\mathbb{E}_n^{*-1}(xx')d\right)^2} .$$

Provided ω^2 exists and is different from zero, $n^{1/2}\omega_n^{-1}(\gamma_n - \gamma) \xrightarrow{d} N(0,1)$, since this is the ratio of two continuously Hadamard differentiable functionals from van der Vaart (1998, Theorem 20.8). Similarly, $n^{1/2}\omega_n^{*-1}(\gamma_n^* - \gamma_n) \xrightarrow{d} N(0,1)$ conditionally on the original sample from Gill (1989). As $\mathrm{RQLR}_n = \left(n^{1/2}\omega_n^{-1}(\gamma_n - \gamma)\right)^2$ and γ is a smooth functional of moments with regular gradients (influence functions), $n^{1/2}\omega_n^{-1}(\gamma_n - \gamma)$ admits an Edgeworth expansion up to $O(n^{-2})$, under moments of order 6 of the gradients and the Cramer condition of Theorem 3.3, see Bhattacharya and Ghosh (1978). Moreover, by symmetry, we have that uniformly in $x \geq 0$

$$\begin{aligned}\Pr\left(n^{1/2}\left|\frac{\gamma_n-\gamma}{\omega_n}\right| \leq x\right) &= \Pr\left(n^{1/2}\frac{\gamma_n-\gamma}{\omega_n} \leq x\right) - \Pr\left(n^{1/2}\frac{\gamma_n-\gamma}{\omega_n} \leq -x\right) \\ &= \Phi(x) - \Phi(-x) \; - \frac{2}{n}\left(\frac{k_2}{2} + \frac{k_4}{24}(x^3-3x)\right. \\ &\left. + \frac{k_6}{72}(x^5 - 10x^3 + 15x)\right)\phi(x) \; + O(n^{-2}) .\end{aligned}$$

Explicit expressions of the cumulants k_2, k_4 and k_6 are given in Bertail and Barbe (1995, Appendix 2), where the coefficients involve both the influence function $\gamma^{(1)}(y,x)$ of γ and the one of ω^2, see also Withers (1983, 1984) and Hall (1992). The bootstrap distribution of $n^{1/2}\omega_n^{*-1}(\gamma_n^* - \gamma_n)$ has the same functional form and thus admits the same Edgeworth expansion with true cumulants replaced by the empirical ones $k_{2,n}$, $k_{4,n}$ and $k_{6,n}$. The result then follows from the fact that $k_{j,n} - k_j = O_P(n^{-1/2})$, for $j = 2, 4, 6$, which is ensured by the moment conditions of order 12 (since k_6 contains moments of order 6 of the gradient, we need moments of order 12 to ensure that a CLT holds). The result about coverage probability follows from Hall (1986). ■

References

Abrevaya, J. (2001): "The Effects of Demographics and Maternal Behavior on the Distribution of Birth Outcomes," *Empir. Econ.*, 26, 247–257.

Aerts, M. and G. Claeskens (2001): "Bootstrap Tests for Misspecified Models, with Application to Clustered Binary Data," *Comput. Statist. Data Anal.*, 36, 383 – 401.

Andersen, P. K. and R. D. Gill (1982): "Cox's Regression Model for Counting Processes: A Large Sample Study," *Ann. Statist.*, 10, 1100–1120.

Andrews, D. W. (2002): "Higher-Order Improvements of a Computationally Attractive k-Step Bootstrap for Extremum Estimators," *Econometrica*, 70, 119–162.

Angrist, J., V. Chernozhukov, and I. Fernández-Val (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica*, 74, 539–563.

Beran, R. (1988): "Prepivoting Test Statistics - A Bootstrap View of Asymptotic Refinements," *J. Am. Stat. Assoc.*, 83, 687–697.

Bertail, P. and P. Barbe (1995): *The Weighted Bootstrap*, vol. 98 of *Lecture Notes in Statistics*, Springer-Verlag New York.

Bhattacharya, R. N. and J. K. Ghosh (1978): "On the Validity of the Formal Edgeworth Expansion," *Ann. Statist.*, 6, 434–451.

Camponovo, L. (2016): "Asymptotic Refinements of Nonparametric Bootstrap for Quasi-Likelihood Ratio Tests for Classes of Extremum Estimators," *Econom. J.*, 19, 33–54.

Chen, X. and D. Pouzo (2009): "Efficient Estimation of Semiparametric Conditional Moment Models with Possibly Nonsmooth Residuals," *J. Econometrics*, 152, 46–60.

Chesher, A. and G. Austin (1991): "The Finite-Sample Distributions of Heteroskedasticity Robust Wald Statistics," *J. Econometrics*, 47, 153–173.

Cribari-Neto, F., S. L. P. Ferrari, and W. A. S. C. Oliveira (2005): "Numerical Evaluation of Tests Based on Different Heteroskedasticity-Consistent Covariance Matrix Estimators," *J. Stat. Comput. Simul.*, 75, 611–628.

Davidson, R. and J. G. MacKinnon (2007): "Improving the Reliability of Bootstrap Tests with the Fast Double Bootstrap," *Comput. Statist. Data Anal.*, 51, 3259–3281.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jacknife," *Ann. Statist.*, 7, 1–26.

Eicker, F. (1967): "Limit Theorems for Regressions with Unequal and Dependent Errors," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, Berkeley, Calif.: University of California Press, 59–82.

Falk, M. and D. Janas (1992): "Edgeworth Expansions for Studentized and Prepivoted Sample Quantiles," *Stat. Probabil. Lett.*, 14, 13–24.

Foutz, R. V. and R. C. Srivastava (1977): "The Performance of the Likelihood Ratio Test When the Model is Incorrect," *Ann. Statist.*, 5, 1183–1194.

Giacomini, R., D. N. Politis, and H. White (2013): "A Warp-Speed Method For Conducting Monte Carlo Experimens Involving Bootstrap Estimators," *Economet. Theor.*, 29, 567–589.

Gill, R. D. (1989): "Non- and Semi-Parametric Maximum Likelihood Estimators and the Von Mises Method (Part 1) [with Discussion and Reply]," *Scand. J. Stat.*, 16, 97–128.

Gine, E. and J. Zinn (1990): "Bootstrapping General Empirical Measures," *Ann. Probab.*, 18, 851–869.

Gonçalves, S. and H. White (2004): "Maximum likelihood and the bootstrap for nonlinear dynamic models," *J. Econometrics*, 119, 199–219.

Gourieroux, C., A. Monfort, and A. Trognon (1984): "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52, 681–700.

Gutenbrunner, C. and J. Jureckova (1992): "Regression Rank Scores and Regression Quantiles," *Ann. Statist.*, 20, 305–330.

Hall, P. (1986): "On the Bootstrap and Confidence Intervals," *Ann. Statist.*, 14, 1431–1452.

——— (1992): *The Bootstrap and Edgeworth Expansion*, Springer Series in Statistics, Springer-Verlag New York.

Hall, P. and J. L. Horowitz (1996): "Bootstrap Critical Values for Tests Based on Generalized-Method-Of-Moments Estimators," *Econometrica*, 64, 891–916.

Hjort, N. L. and D. Pollard (1993): "Asymptotics for Minimisers of Convex Processes," Tech. rep., University of Oslo.

Huber, P. J. (1967): "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, Berkeley, Calif.: University of California Press, 221–233.

Kent, J. T. (1982): "Robust Properties of Likelihood Ratio Tests," *Biometrika*, 69, 19–27.

Kim, J. H. (2003): "Higher-Order Improvements of the Restricted Parametric Bootstrap for Tests," National University of Singapore.

Kim, T.-H. and H. White (2003): "Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regression," in *Advances in Econometrics*, Bingley: Emerald (MCB UP), vol. 17, 107–132.

Kline, P. and A. Santos (2012): "Higher Order Properties of the Wild Bootstrap Under Misspecification," *J. Econometrics*, 171, 54–70.

Kocherginsky, M., X. He, and Y. Mu (2005): "Practical Confidence Intervals for Regression Quantiles," *J. Comput. Graph. Statist.*, 14, 41–55.

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

Lee, S. M. S. and P. Yang (2020): "Bootstrap Confidence Regions Based on M-Estimators Under Nonstandard Conditions," *Ann. Statist.*, 48, 274–299.

Lien, D. and Q. H. Vuong (1987): "Selecting the Best Linear Regression Model: A Classical Approach," *J. Econometrics*, 35, 3 – 23.

Long, J. S. and L. H. Ervin (2000): "Using heteroscedasticity consistent standard errors in the linear regression model," *Am. Stat.*, 54, 217–224.

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

Maistre, S., P. Lavergne, and V. Patilea (2017): "Powerful Nonparametric Checks for Quantile Regression," *J. Stat. Plan. Infer.*, 180, 13 – 29.

Mammen, E. (1992): "Bootstrap, Wild Bootstrap, and Asymptotic Normality," *Probab. Theory Rel.*, 93, 439–455.

Marcellino, M. and B. Rossi (2008): "Model Selection for Nested and Overlapping Nonlinear, Dynamic and Possibly Mis-specified Models," *Oxford B. Econ. Stat.*, 70, 867–893.

Powell, J. (1991): "Estimation of Monotonic Regression Models UUnder Quantile Restrictions," in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, ed. by W. A. Barnett, J. Powell, and G. E. Tauchen, Cambridge University Press.

Royall, R. M. (1986): "Model Robust Confidence Intervals Using Maximum Likelihood Estimators," *Int. Stat. Rev.*, 54, 221.

Spokoiny, V. and M. Zhilova (2015): "Bootstrap Confidence Sets under Model Misspecification," *Ann. Statist.*, 43, 2653–2675.

Stafford, J. E. (1996): "A Robust Adjustment of the Profile Likelihood," *Ann. Statist.*, 24, 336–352.

Tarr, G. (2012): "Small Sample Performance of Quantile Regression Confidence Intervals," *J. Stat. Comput. Simul.*, 82, 81–94.

van der Vaart, A. W. (1998): *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307–333.

White, H. (1980): "Using Least Squares to Approximate Unknown Regression Functions," *Int. Econ. Rev.*, 21, 149–70.

——— (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, pp. 1–25.

Withers, C. S. (1983): "Expansions for the Distribution and Quantiles of a Regular Functional of the Empirical Distribution with Applications to Nonparametric Confidence Intervals," *Ann. Statist.*, 11, 577–587.

——— (1984): "Asymptotic Expansions for Distributions and Quantiles with Power Series Cumulants," *J. Roy. Stat. Soc. B.*, 46, 389–396.

Wu, C. F. J. (1986): "Jackknife, Bootstrap and Other Resampling Methods in Regression-Analysis - with Discussion," *Ann. Statist.*, 14, 1261–1295.