Article

Robust Variable Selection and Regularization in Quantile Regression Based on Adaptive-LASSO and Adaptive E-NET

Innocent Mudhombo and Edmore Ranganai

Department of Statistics, University of South Africa, Florida Campus, Private Bag X6, Florida Park, Roodepoort 1710, South Africa; 31334334@mylife.unisa.ac.za
* Correspondence: rangae@unisa.ac.za; Tel.: +27-(0)11-6709257

Abstract: Although the variable selection and regularization procedures have been extensively considered in the literature for the quantile regression (QR) scenario via penalization, many such procedures fail to deal with data aberrations in the design space, namely, high leverage points (X-space outliers) and collinearity challenges simultaneously. Some high leverage points referred to as collinearity influential observations tend to adversely alter the eigenstructure of the design matrix by inducing or masking collinearity. Therefore, in the literature, it is recommended that the problems of collinearity and high leverage points should be dealt with simultaneously. In this article, we suggest adaptive LASSO and adaptive E-NET penalized QR (QR-ALASSO and QR-AE-NET) procedures where the weights are based on a QR estimator as remedies. We extend this methodology to their penalized weighted QR versions of WQR-LASSO, WQR-E-NET procedures we had suggested earlier. In the literature, adaptive weights are based on the RIDGE regression (RR) parameter estimator. Although the use of this estimator may be plausible at the \( \tau = 0.5 \) for the symmetrical distribution, it may not be so at extreme quantile levels. Therefore, we use a QR-based estimator to derive adaptive weights. We carried out a comparative study of QR-LASSO, QR-E-NET, and the ones we suggest here, viz., QR-ALASSO, QR-AE-NET, weighted QR LASSO penalized and weighted QR adaptive AE-NET penalized (WQR-LASSO and WQR-AE-NET) procedures. The simulation study results show that QR-ALASSO, QR-AE-NET, WQR-LASSO and WQR-AE-NET generally outperform their nonadaptive counterparts. At predictor matrices with collinearity inducing points under normality, the QR-ALASSO and QR-AE-NET, respectively, outperform the non-adaptive procedures in the unweighted scenarios, as follows: in all 16 cases (100%) with respect to correctly selected (shrunken) zero coefficients; in 88% with respect to correctly fitted models; and in 81% with respect to prediction. In the weighted penalized WQR scenarios, WQR-LASSO and WQR-AE-NET outperform their non-adaptive versions as follows: in 75% of the time with respect to both correctly fitted models and correctly shrunken zero coefficients and in 63% with respect to prediction. At predictor matrices with collinearity masking points under normality, the QR-ALASSO and QR-AE-NET, respectively, outperform the non-adaptive procedures in the unweighted scenarios as follows: in prediction, in 100% and 88% of the time; with respect to correctly fitted models in 100% and 50% (while in 50% equally); and with respect to correctly shrunken zero coefficients in 100% of the time. In the weighted scenario, WQR-LASSO and WQR-AE-NET outperform their respective non-adaptive versions as follows; with respect to prediction, both in 63% of the time; with respect to correctly fitted models, in 88% of the time while with respect to correctly shrunken zero coefficients in 100% of the time. At predictor matrices with collinearity inducing points under the \( t \)-distribution, the QR-ALASSO and QR-AE-NET procedures outperform their respective non-adaptive procedures in the unweighted scenarios as follows: in prediction, in 100% and 75% of the time; with respect to correctly fitted models 88% of the time each; and with respect to correctly shrunken zero 88% and in 100% of the time. Additionally, the procedures WQR-LASSO and WQR-AE-NET and their unweighted versions result in the former outperforming the latter in all respective cases with respect to prediction whilst there is no clear “winner” with respect to the other two measures. Overall, the WQR-LASSO generally outperforms all other models with respect to all measures. At the predictor matrix with collinearity-masking points under the \( t \)-distribution, all adaptive versions outperformed their respective non-adaptive versions with respect to all metrics.
the unweighted scenarios, the QR-ALASSO and QR-AE-NET dominate their non-adaptive versions as follows: in prediction, in 63% and 75% of the time; with respect to correctly fitted models, in 100% and 38% (while in 62% equally); in 100% of the time with respect to correctly shrunk zero coefficients. In the weighted scenarios, all adaptive versions outperformed their non-adaptive versions as follows: 62% of the time in both respective cases with respect to prediction while it is vice-versa with respect to correctly fitted models and with respect to correctly shrunk zero coefficients. In the weighted scenarios, WQR-ALASSO and WQR-AE-NET dominate their respective non-adaptive versions as follows; with respect to correctly fitted models, in 62% of the time while with respect to correctly shrunk zero coefficients in 100% of the time in both cases. At the design matrix with both collinearity and high leverage points under the heavy-tailed distributions (t-distributions with \(d \in (1; 6)\) degrees of freedom) scenarios, the dominance of the adaptive procedures over the non-adaptive ones is again evident. In the unweighted scenarios, the procedures QR-ALASSO and QR-AE-NET outperform their non-adaptive versions as follows; in prediction, in 75% and 62% of the time; with respect to correctly fitted models, they perform better in 100% and 88% of the time, while with respect to correctly shrunk zero coefficients, they outperform their non-adaptive ones 100% of the time in both cases. In the weighted scenarios, WQR-ALASSO and WQR-AE-NET dominate their non-adaptive versions as follows; with respect to prediction, in 100% of the time in both cases; and with respect to both correctly fitted models and correctly shrunk zero coefficients, they both do 88% of the time. Results from applications of the suggested procedures to real life data sets are more or less in line with the simulation studies results.

**Keywords:** weighted quantile regression; adaptive LASSO penalty; penalty; adaptive E-NET penalty; collinearity inducing point; collinearity hiding point; collinearity influential points

1. **Introduction**

Variable selection is at the heart of the model-building process. However, it is fraught with challenges. These challenges stem from data aberrations in both the X-space and Y-space. X-space data aberrations comprise extreme observations referred to as high leverage points, while Y-space extreme ones are residual outliers referred to as outliers. High leverage points can be collinearity hiding or inducing points, referred to as collinearity influential observations [1]. The least squares (LS) method is susceptible to both X-space and Y-space data aberrations (non-Gaussian error terms), hence the need for robust procedures. In the literature, the least absolute deviation (LAD)-based [2] procedures have been suggested as alternatives to the LS as they are robust in the Y-space. One such LAD-based procedure is quantile regression (QR) [3]. However, LAD-based procedures are amenable to high leverage points. QR, which generalizes the LAD estimator to all quantile levels, is a very attractive procedure, as it provides more information about the conditional distribution of \(Y\) given \(X\). As a result, it has generated more research interest in recent years, hence the need for robust variable selection procedures in the QR framework.

Subset selection has been a topical issue since the 19th century. However, subset selection procedures tend to be unstable and especially unsuitable for variable selection when the dimension is high [4]. As a result, towards the end of the 20th century, the literature witnessed a fair proliferation of penalization (shrinkage) procedures as the alternative variable/model selection and regularization tools. Penalization procedures tend to be fairly stable and produce lower prediction errors than subset selection procedures. So they have been suggested to proffer solutions to shortcomings of subset selection methods, both in LS and QR scenarios, with varying degrees of success. These methods include procedures such as RIDGE [5], LASSO [6], elastic net (E-NET) [7], the nonnegative garotte selection [4,8] and their extended versions, just to mention a few.

Although the LASSO penalty has the edge over the RIDGE penalty against overfitting (as the RIDGE penalty does not shrink the coefficient estimates to zero) with attractive
prediction properties, its model selection criterion is only consistent under some restrictive assumptions [9]. Other drawbacks of the LASSO are that in the $p > n$ case, where $p$ and $n$ are the numbers of predictors and observations, respectively, the LASSO selects at most $n$ variables before it saturates; in the case where a group of variables among which the pairwise correlations are very high, then the LASSO tends to select only one variable from the group randomly; for the usual $n > p$ case with high correlated predictors, it has been empirically observed that the prediction performance of the LASSO is dominated by RIDGE regression [7]. Zou and Hastie [7] suggested the E-NET procedure with similar sparsity of representation to the LASSO often outperforming the LASSO procedure. The competitive edge of the E-NET over the LASSO is that it encourages a grouping effect, where strongly correlated predictors tend to be in or out of the model together; and is particularly useful when $p$ is much bigger than $n$.

The LASSO penalty performs poorly in the presence of high leverage points (especially collinearity influential points), and procedures such as smoothly clipped absolute deviation (SCAD) [10] and minimax concave penalty (MCP) [11] are better options. Weights-based regression methods to remedy the influence of both high leverage points and outliers have been suggested in the literature. In the LS scenario, weighted regression has been used in variable selection, for example, doubly adaptive penalized regression, which satisfies both sparse and robustness properties [12]. In the QR scenario, weighted QR (WQR) has been suggested as a remedy to high leverage points [13]. Building upon this idea, Ranganai and Mudhombo [14] suggested a weighted QR LASSO (WQR-LASSO) variable selection and regularization parameter estimation procedure, which is robust to both the $X$ and $Y$-spaces data aberrations.

Since LASSO penalizes the parameter coefficient estimates equally and is only consistent under some restrictive assumptions, a dynamic variable selection and regularization approach has been suggested via the adaptive LASSO (ALASSO) penalty. Compared to LASSO, ALASSO enjoys the oracle properties which guarantee optimal performance in large samples, with high dimensionality [15] as well as the computational advantage of the LASSO owing to the efficient path algorithm [16]. The ALASSO procedure relies on suitably chosen weights. Frommlet and Nuel [17] suggested the adaptive RIDGE (ARIDGE) procedure for variable selection with interesting formulations of the adaptive weight. In the QR scenario, variable selection has been suggested via ALASSO QR (QR-ALASSO) [18]. Zou and Zhang [19] suggested the adaptive E-NET (AE-NET) procedure that inherits some good properties from both ALASSO and A RIDGE penalty-based procedures.

This article suggests adaptive LASSO and adaptive E-NET penalized weighted QR (WQR-ALASSO and WQR-AE-NET) procedures with the initial coefficient estimator used to compute adaptive weights need not be consistent. This coefficient estimate is computed from WQR in which the weights used to downweigh high leverage points are based on the minimum covariance determinant (MCD) estimator of Rousseeuw [20]. The WQR-ALASSO and WQR-AE-NET procedures are extensions of Ranganai and Mudhombo [14]'s WQR-LASSO and WQR-E-NET procedures as mitigation against both high leverage and collinearity influential observations.

In summary, the motivations of this study are premised on the following:

- Rather than carrying an “omnibus” study of adaptive penalized QR, we carry out a detailed study by distinguishing different types of high leverage points under different distribution scenarios viz.
  - Collinearity influential points that comprise collinearity inducing and collinearity masking ones.
  - A “mixture” of collinearity and high leverage points that are not collinearity influential.
- Unlike the conditional mean regression (LS) estimator, which is a global one, the regression quantile (RQ) estimator is a local one. Therefore, we suggest a QR-based estimator instead of (RR) parameter-based estimator suggested in the literature to
derive adaptive weights in extending QR-LASSO and QR-E-NET procedures to QR-
ALASSO and QR-AE-NET procedures, respectively.
• We further extend QR-ALASSO and QR-AE-NET procedures to WQR-ALASSO and
WQR-AE-NET procedures using the same methodology.
• We carry a comparative study of these models using simulation studies and well-
known data sets from the literature.

The rest of the paper is organized as follows: In Section 2.1, a brief overview of QR, the
adaptive weights and the QR-ALASSO and QR-AE-NET procedures are given. The QR-
ALASSO and QR-AE-NET procedures are extended to their WQR-ALASSO and WQR-
AE-NET counterparts, respectively, in Section 2.2. Simulation results and examples are
presented in Section 3. The simulations consider collinearity influential observations in the
design matrix as well as normal and t-distributed error term scenarios with varying degrees
of tail heaviness. The examples consider data sets from the literature with collinearity
influential observations. Finally, concluding remarks are provided in Section 4.

2. Quantile Regression

Consider the linear regression model given by
\[ y_i = x_i' \beta + \epsilon_i, \quad i \in [1 : n], \] (1)
where \( y_i \) is the \( i \)th response random variable, \( x_i \) the \( i \)th row of design matrix \( X \in \mathbb{R}^{n \times p} \), \( \beta \) is vector of parameters and \( \epsilon_i \sim F \), the \( i \)th error term. The QR estimator [3] for the
coefficient vector \( \beta \in \mathbb{R}^p \) in Equation (1), is based on an optimization problem solved
via linear programming techniques. The QR minimization problem for estimating the
parameter vector is given by
\[ \hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i' \beta(\tau)), \quad i \in [1 : n], \] (2)
with the check function,
\[ \rho_\tau(u) = \begin{cases} \tau u, & \text{if } u \geq 0 \\ (\tau - 1)u, & \text{if } u < 0 \end{cases} \quad i \in [1 : n] \]
denoting the re-weighting function of residuals \( \epsilon_i \)'s for \( 0 < \tau < 1 \), where
\( u = y_i - x_i' \beta(\tau) \)
denotes residuals and \( \hat{\beta}(\tau) \) the \( \tau \)th RQ.

2.1. Variable Selection in Quantile Regression

In this section, we give an overview of QR variable selection using penalized proce-
dures. We discuss QR-penalized methods with LASSO [6] and E-NET [7] based penalties
as well as the ALASSO [16] and AE-NET [19] based penalties.

The LASSO-penalized QR procedure, denoted by QR-LASSO, is a QR variable selec-
tion procedure that uses a LASSO (\( \ell_1 \)) penalty. This penalized QR procedure is given by the
minimization problem
\[ \hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i' \beta(\tau)) + n\lambda \sum_{j=1}^{p} |\beta_j|, \quad j \in [1 : p], \quad i \in [1 : n], \] (3)
where the tuning parameter \( \lambda \) shrinks the \( \beta \) coefficients towards zero equally. The second
term represents the penalty term and other terms are explained in Equations (1) and (2). The
continuous shrinkage by QR-LASSO procedure frequently enhances prediction precision
due to bias–variance trade-off.
The QR variable selection procedure that uses an ALASSO penalty is denoted by QR-ALASSO. The adaptive weights penalize different coefficients differently in the QR setting, thereby outperforming the LASSO penalty scenario. The QR-ALASSO procedure is an extension of the LASSO penalty-based method. The construction of adaptive weights is discussed in Section 2.1.1. The QR-ALASSO procedure selects variables by solving a minimization problem:

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_i [y_i - x_i' \beta(\tau)] + \lambda \sum_{j=1}^{p} \hat{\omega}_j |\beta_j|, \text{ for } j \in [1 : p], \ i \in [1 : n],$$  \hspace{1cm} (4)

where $\hat{\omega}_j$ is the adaptive weight and the tuning parameter $\lambda_j = \hat{\omega}_j \lambda$, shrinks predictors coefficients to zero differently. Other terms are as defined in Equations (1) and (2). The QR-ALASSO tuning parameter is no-longer constant ($\lambda$), but $\lambda_j$ for $j = 1, 2, \ldots, p$, which is varying.

We present the penalized QR procedure by [7]'s E-NET penalty (QR-E-NET). This penalized QR procedure comprises a combination of the LASSO and RIDGE penalties. The E-NET-penalized QR is best applicable to unidentified groups of predictors. The E-NET-penalized QR is given by the minimization problem

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_i [y_i - x_i' \beta(\tau)] + \alpha \lambda \sum_{j=1}^{p} |\beta_j| + (1 - \alpha) \lambda \sum_{j=1}^{p} \hat{\omega}_j^2 |\beta_j|, \ j \in [1 : p], \ i \in [1 : n],$$  \hspace{1cm} (5)

where $\alpha \in [0, 1]$ is the mixing parameter between RIDGE ($\alpha = 0$) and LASSO ($\alpha = 1$), and $\lambda$ is the tuning parameter for the second and third terms. This procedure outperforms its RIDGE and LASSO counterparts [7].

In a similar fashion to the extension of QR-LASSO (Equation (3)) to QR-ALASSO (Equation (4)), QR-E-NET (Equation (5)) can be extended to QR-AE-NET (Equation (6)) as in [19]. Suppose the adaptive weight ($\hat{\omega}_j$) is constructed as in Equation (8). We define the QR-AE-NET estimator $\hat{\beta}$ as

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_i [y_i - x_i' \beta(\tau)] + \alpha \lambda \sum_{j=1}^{p} \hat{\omega}_j |\beta_j| + (1 - \alpha) \lambda \sum_{j=1}^{p} \hat{\omega}_j \beta_j^2, \ j \in [1 : p], \ i \in [1 : n],$$  \hspace{1cm} (6)

where the terms are defined as in Equations (4) and (5). QR-AE-NET reduces to QR-ALASSO for $\alpha = 1$ and QR-ARIDGE for $\alpha = 0$. QR-AE-NET inherits the desired optimal minimax bound from ALASSO [16]. It is further anticipated that the procedure deals with collinearity challenges better due to the presence of the $\ell_2$-penalty.

2.1.1. Choice of Adaptive Weights for ALASSO

Consider the penalty construction $\sum_{j=1}^{p} |\hat{\beta}_j|^q$ as a general penalty term in most penalization problems. The adaptive penalty can be expressed as $\sum_{j=1}^{p} \omega_j |\hat{\beta}_j|^q$, where $\omega_j$s are the adaptive weights. When $q = 1$, we obtain an ALASSO penalty and for $q = 2$, we obtain an adaptive RIDGE (ARIDGE) penalty. [16] suggested using the LS $\hat{\beta}$ estimator in determining adaptive weights, $\omega_j$s. The LS $\hat{\beta}$ is suitable in the absence of collinearity. In the presence of collinearity, Ref. [16] suggested the use of RIDGE $\hat{\beta}$ as a suitable replacement because it is superior in stability than its LS counterpart. In constructing the adaptive weights ($\omega_j$) in the LS case, we suggest using the MCD-based weighted RIDGE regression (WRR) estimator

$$\hat{\beta}_{WRR} = \arg\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \omega_i (y_i - x_i' \beta)^2 + n \lambda \sum_{j=1}^{p} \beta_j^2, \ j \in [1 : p], \ i \in [1 : n],$$  \hspace{1cm} (7)
where $\lambda$ is the tuning parameter, $\beta$ is a vector of parameters and $\tilde{\omega}_i$ is a robust MCD-based weight [14]. Thus, in the $LS$ case the weights are given by

$$\omega_j = \left( |\hat{\beta}_{WRR_j}| + 1/n \right)^{-\gamma},$$

(8)

where $\hat{\beta}_{WRR_j}$ is $j$th parameter estimator from a WRR penalized solution and $1/n$ is added to avoid dividing by zero. This adaptive weight is a special case of weight, $\omega = (|\hat{\beta}_{RR_j}|^\gamma + \delta^\gamma)^{(1/\gamma)/2}/(\gamma \tau)$ proposed by [17], where $\theta = 1, \delta = 1/n$ and $\gamma = 1$.

Although the use of $\omega_i$ may be applicable to the $\ell_1$ estimator ($RQ$ at $\tau = 0.5$) for the symmetrical distribution, it may not be applicable at extreme quantile levels in the presence of high leverage (and collinearity influential) points. This is due to the fact that the presence of these atypical observations result in $RQ$ planes often crossing (unequal slope parameter estimates) (see [21,22]), although they are theoretically parallel. Therefore, instead of using $\hat{\omega}_{WRR_j}$, we suggest using the one based on $WQR$, $\tilde{\beta}(\tau)_{WQR}$, which is the solution to the minimization problem

$$\hat{\beta}(\tau)_{WQR} = \arg \min_{\beta \in \mathcal{R}^p} \sum_{i=1}^n \tilde{\omega}_i |y_i - x'_i \beta| + n \lambda \sum_{j=1}^p \beta_j^2, j \in [1 : p], \ i \in [1 : n],$$

(9)

where $\hat{\beta}(\tau)_{WQR}$ is the $\tau$th $WQR$-RIDGE estimator, and $\lambda$ and $\tilde{\omega}_i$ are as defined in Equations (2) and (7). An analogous result to Equation 8 ($\omega_i$) based on using the $WQR$-RIDGE parameter estimates $\tilde{\beta}(\tau)_{WQR}$ becomes

$$\tilde{\omega}_j = \left( |\tilde{\beta}(\tau)_{WQR_j}| + 1/n \right)^{-1},$$

(10)

where $\tilde{\beta}(\tau)_{WQR_j}$ is the $j$th parameter estimator from a $WQR$-RIDGE penalized solution at a specified $\tau$ quantile level, and other terms are defined in Equation (8). Therefore, for all variable selection in simulations and applications, we use the weights $\tilde{\omega}_j$ for $WQR$ instead of $\omega_j$ for $WRR$ ($\tilde{\omega}_j$ is same as $\omega_j$ at $\tau = 0.50$ $RQ$ level). $\hat{\beta}(\tau)_{WQR}$ is more representative at all $\tau$ quantiles than $\hat{\beta}_{WRR_j}$. To our knowledge, no such adaptive weight has been applied in a $Q$R variable selection scenario. The adaptive weight has the advantage of not being influenced by extreme points and adjustable to particular distribution levels, for example, $t_1$, $t_2$ distributions, et cetera and applicable at all $\tau$ quantile levels.

2.2. Adaptive Penalized Weighted Quantile Regression

Using the adaptive weights $\tilde{\omega}_j$ and the MCD-based weights $\omega_j$, we propose an adaptive penalized $WQR$ variable selection procedure. The $WQR$ penalization problems culminate from $ALASSO$ and $AE-NET$ penalties, namely the $WQR$ adaptive LASSO ($WQR$-ALASSO) and the $WQR$ adaptive E-NET ($WQR$-AE-NET) estimators, respectively. The approach is robust in the presence of high leverage points (and collinearity influential observations) due to robustly chosen MCD-based weights as in the $WQR$ (see [14,23,24]).

We first write the proposed $WQR$-ALASSO given by the minimization problem

$$\hat{\beta}(\tau) = \arg \min_{\beta \in \mathcal{R}^p} \sum_{i=1}^n \tilde{\omega}_i |y_i - x'_i \beta(\tau)| + \lambda \sum_{j=1}^p \tilde{\omega}_j |\beta_j|, \ for \ j \in [1 : p], \ i \in [1 : n],$$

(11)

where $\tilde{\omega}_i$s are $WQR$-RIDGE parameter estimator-based adaptive weights, $\tilde{\omega}_i$s are MCD-based weights and, other terms are as defined in Equations (2) and (3). The proposed adaptive penalized $QR$ procedure, $WQR$-AE-NET is the minimization problem

$$\hat{\beta}(\tau) = \arg \min_{\beta \in \mathcal{R}^p} \sum_{i=1}^n \tilde{\omega}_i |y_i - x'_i \beta(\tau)| + \alpha \lambda \sum_{j=1}^p \tilde{\omega}_j |\beta_j| + (1 - \alpha) \lambda \sum_{j=1}^p \tilde{\omega}_j \beta_j^2, j \in [1 : p], \ i \in [1 : n],$$

(12)
where the terms are defined in Equations (2), (9) and (11). Special cases of this construction are WQR-ALASSO if $\alpha = 1$ and WQR-ARIDGE if $\alpha = 0$. Just like its unweighted QR counterpart, WQR-AE-NET inherits the desired optimal minimax bound from ALASSO [16]. The next section discusses the asymptotic properties of our proposed procedures.

2.3. Asymptotic Properties

In this section, we establish the oracle properties of the adaptive LASSO penalized QR. Consider the linear model in Equation 1 and that $P(\varepsilon_i < 0) = \tau$. Let us set $X_1 = (x_{11}, \ldots, x_{n_0}) \in R^{n_0}$ and set $X_2 = (x_{n_0+1}, \ldots, x_p) \in R^{p-n_0}$ (noise variables) such that $\beta = (\beta_1', \beta_2')'$, where the true regression coefficient vector $\beta_1$ corresponds to non-zero coefficients and $\beta_2 = 0$.

For asymptotic normality, the following theoretical results are assumed to be true for a suitable choice of $\lambda_n$ [18]:

(i) The regression errors $\varepsilon_i$'s, are i.i.d., with $\tau$th quantile zero and a continuous, positive density $f(.)$ in a neighborhood of zero and $F$ distributed [25]. NOTE: $F(0) = \tau$ and $|f(y) - f(0)| \leq c|y|^{1/2}$, $\forall y$ in the neighborhood of 0 and real quantile $\tau \in (0, 1)$.

(ii) Let $\Omega = diag(\omega_1, \omega_2, \ldots, \omega_n)$, where $\omega_i$ for $i = 1, 2, \ldots, n$ are known positive values that satisfy $\max(\omega_i) = O(1)$ [14].

(iii) Consider the design $X$, such that there exists a positive definite matrix $\Sigma$, where $\Sigma = \lim_{n \to \infty} X'X/n$ with

$$X'X = \Sigma_{11} \Sigma_{12} \Sigma_{21} \Sigma_{22},$$

where $\Sigma_{11} \in R^{n_0}$ and $\Sigma_{22} \in R^{p-n_0}$.

Let Equation 6 have $\hat{\beta}^{ALASSO}$ as its solution. We now state a theorem for an asymptotic oracle property for i.i.d. random error terms (Theorem 1).

**Theorem 1.** Consider a sample $\{(x_i, y_i), i = 1, \ldots, n\}$ from Equation (6) satisfying Conditions (i) and (iii) with $\Omega = I_n$ (constant weight of 1). If $\sqrt{n}\lambda_n \to 0$ and $n^{(\gamma + 1)/2}\lambda_n \to \infty$, then we have

1. Sparsity: $\hat{\beta}^{ALASSO}_2 = 0$;
2. $\sqrt{n}(\hat{\beta}_1^{ALASSO} - \beta_1^{ALASSO})$ converges asymptotically in limit to $N\left(0, \frac{\tau(1-\tau)\Sigma_{11}}{f(0)^2}\right)$.

Now, consider the extension of oracle results to non-i.i.d. random error scenarios. The following assumptions are considered.

(iv) As $n \to \infty$, $\max_{i < n} \{x_i'x_i/n\} \to 0$.

(v) The random errors $\varepsilon_i$s are independent with $F_i(t) = P(\varepsilon_i \leq t)$ the distribution function of $\varepsilon_i$. We assume that each $f_i(.)$ is locally linear near zero (with a positive slope) and $F_i(0) = \tau$. Define $\Psi_n(t) = \int_0^t \sqrt{n}F_i(s)\sqrt{n} - F_i(0)ds$, which is a convex function for each $n$ and $i$.

(vi) Assume that, for each $u_i$, $(1/n)\sum_{i=1}^n \Psi_n(u'x) \to \zeta(u)$, where $\zeta(.)$ is a strictly convex function taking values in $[0, \infty)$.

**Corollary 1.** Under Conditions (ii), (iii) and (iv), Theorem 1 holds, provided the non-i.i.d. random errors satisfy (v) and (vi) (see also [26]).

**Remark 1.** The QR model, a non-i.i.d. random error model, is catered for by (v) [27]. See the proofs online [18].

3. Simulation Study

In this Section, we discuss the simulation results of four adaptive variable selection and estimation procedures, namely QR-ALASSO, QR-AE-NET (for $\alpha = 0.5$), WQR-ALASSO,
and WQR-AE-NET (for $\alpha = 0.5$) against their non-adaptive versions. We evaluate and illustrate these four procedures’ variable selection and prediction performance under normality and $t$-distributions (heavy-tailed error distributions) in the presence of collinearity influential and high leverage observations in the design space. The pairwise comparisons are two-pronged, viz., (i) comparison of adaptive (QR-LASSO and QR-AE-NET) versus non-adaptive (QR-LASSO and QR-E-NET) procedures and (ii) comparison of weighted adaptive (WQR-LASSO and WQR-AE-NET) procedures versus weighted non-adaptive (WQR-LASSO, WQR-E-NET) procedures. Cross comparisons are explored between (i) and (ii) scenarios. All simulations are applied at two QR levels, $\tau \in (0.25; 0.50)$ and two sample sizes, $n \in (50; 100)$, although $n = 100$ is left for brevity out in the presentation of the results for brevity.

3.1. Simulation Design Scenarios

In the design space, we consider four collinearity influential point scenarios in addition to the orthogonal design matrix and a correlated design matrix with high leverage points, while in the error term distribution scenarios, we consider the normal and heavy tailed ($t$-distribution cases at different degrees of freedom) distribution cases. The design matrices choices are simulated as in Jongh et al. [14,28,29], viz.

(i) $D1$—the well-behaved orthogonalized $n \times p$ design matrix $X$ (where the initial un-orthogonalized’s $p = 8$ columns are generated from $N(0,1)$) satisfy the condition $X'X = nI$. We first generate the $n \times p$ response data, $W$, where $w_{ij} \sim N(0,1)$ with $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,p$. We find the singular value decomposition (SVD) of the design matrix $W$, given by $W = UDV'$, where $U$ and $V$ are orthogonal with the diagonal entries of $D$. The diagonal entries of $D$ are the eigenvalues of the design matrix $W$. Finally, the design matrix $X$ is given by $X = \sqrt{n}U$ such that $X'X = nI$ since $U$ is orthogonal. We use the design matrix $D1$ as a baseline when comparing with scenarios $D2$–$D5$.

(ii) $D2/D4$—the design matrix $D1$ with the most extreme point by Euclidean distance moved 10 units in the $X$ direction ($D2$) and 100 units in the $X$ direction ($D4$). The resultant extreme points are collinearity inducing points for scenarios $D2$ and $D4$ (see Figure 1).

(iii) $D3/D5$—the design matrix $D1$ with the most and second most extreme points by Euclidean distance moved 10 and 100 units, respectively, in the $X$ direction. The two extreme points have a masking effect on collinearity (see scenarios $D3$ and $D5$ in Figure 1).

$D6$—a correlated design matrix case with high leverage points [14,30]. In $D6$, we partitioned the $n \times p$ design matrix, $X = (X_1', X_2')'$, where the uncontaminated part, $X_1 \sim N(0, V)$ and the contaminated $m \times p$ sub-matrix $X_2 \sim N(1, I)$ ($I$ is the mean vector of ones and $I$ is an identity matrix). The exponential decay $0.5^{j-i}$, for $i = 1,2,3,4,5,6,7,8$, $j = 1,2,3,4,5,6,7,8$ generates the $(ij)$th entry of covariance matrix $V$ and 0 is the mean vector of zeros. The design matrix $D6$ has $m = (5; 10)$ contamination points (using contamination rate of 10%) from $n = (50; 100)$ observations.

In the Gaussian error term case, for $n$ observations and $p$ predictor variables, the dependent variable is determined by

$$Y_i = X_i'\beta + \sigma e_i, \text{ for } i \in [1:n].$$

(13)

where $\sigma \in (1;3)$, is standard deviation, $e_i \sim N(\mu, \sigma^2)$ is the error term, $\beta = (3,1.5,0,0,2,0,0,0,0)$ and other terms are defined in Equation 1 ($D1$–$D5$). The dependent variable in heavy tailed distributions is determined by Equation 1 with standard error $\sigma \in (0.5;1)$ and $d \in (1;6)$ degrees of freedom ($e_i \sim t_d$) ($D6$).
The number of simulation runs in each of the D1-D6 scenarios is 200, and 10-fold cross-validation (CV) is applied to select the best tuning parameter (optimal $\lambda$) in the test data.

Different scenarios of the ALASSO or AE-NET penalized QRs (weighted and unweighted), are explored using hqreg package, a readily available R-add-on-package [31]. The hqreg package, a semi-smooth Newton coordinate descent algorithm, chooses optimal $\lambda$ as in Ranganai and Mudhombo [14].

Figure 1. Collinearity influential points. First row panel shows boxplots for collinearity inducing points in D2 and D4, and second row panel shows boxplots for collinearity masking points (collinearity hiding points) in D3 and D5.

3.2. Results

To study the robustness and performance of the different ALASSO and AE-NET-penalized QR procedures in the presence of collinearity influential points, high leverage points and influential points under different distributions, we present the simulated results. These results are based on the metrics; the median of the test errors, $\frac{\text{median}}{\text{median}} \{ \epsilon_i \}$ and its respective measure of dispersion MAD of the test errors, percentage (%) correctly fitted, average of correct/incorrect zero coefficients. The MAD of test errors for each penalized estimator is estimated by $\text{MAD} = 1.4826 \left( \frac{\text{median}}{\text{median}} \{ \epsilon_i \} \right)$. We compare the performance of non-adaptive procedures (QR-LASSO and QR-E-NET) and the adaptive ones we suggest here based on a QR adaptive weight, viz., QR-ALASSO, QR-AE-NET, WQR-ALASSO and WQR-AE-NET procedures. Without loss of generality, we only consider all procedures at $\tau = 0.25$ and $\tau = 0.50$ QR levels.
Remark 2. The five zero coefficients correspond to the set \( \{ \beta_3, \beta_4, \beta_6, \beta_7, \text{ and } \beta_8 \} \), hence the maximum average of correctly/incorrectly selected (shrunk) coefficients is 5, while the set of correctly selected models is given as a proportion, i.e., a percentage.

3.2.1. D1 under the t-Distribution

Results at the design matrix D1 (the orthogonal design scenario or baseline case) under normality and heavy-tailed t-distributed error scenarios with degrees of freedom \( d = 1 \) are shown in Table 1. In all 16 pairwise cases except one, QR-ALASSO and QR-AE-NET procedures outperform their respective non-adaptive penalized procedures at \( \tau \in (0.25; 0.50) \) QR levels with respect to all measures except one. The exception is the LASSO-penalized procedure, which outperforms all adaptive penalized procedures in the heavy-tailed distribution scenario (t-distribution case) when \( d = 1, \sigma = 1 \) and \( \tau = 0.50 \). The QR-ALASSO outperforms the QR-AE-NET procedure in both the Gaussian and t-distribution cases in this baseline scenario with respect to MAD of the Test Errors. Generally, the adaptive penalized procedures, QR-ALASSO and QR-AE-NET correctly shrinks covariate parameters \( \{ \beta_j; j = 3, 4, 6, 7, 8 \} \) to zero with high precision (in all cases) compared to their respective non-adaptive penalized procedures. However, the performance with respect to all measures decreases as the error term distribution becomes heavier.

### Table 1. Penalized unweighted quantile regression at D1 for \( n = 50 \) (\( \tau = 0.25 \) and \( \tau = 0.50 \)) under the normal and heavy tailed t-distributions; bold text indicate superior performance.

| Distribution | Parameter | Method       | Median(MAD) Test Error | Correctly Fitted | No of Zeros | Median(\( \lambda \)) |
|--------------|-----------|--------------|------------------------|------------------|------------|----------------------|
| D1: N-distribution | \( \sigma = 1, \tau = 0.25 \) | QR-LASSO | 0.72(1.17) | 62.00 | 4.40 | 0.00 | 0.04 |
| | | QR-E-NET | 0.75(1.26) | 16.50 | 3.35 | 0.00 | 0.03 |
| | | QR-ALASSO | 0.71(1.10) | 99.50 | 5.00 | 0.00 | 0.03 |
| | \( \sigma = 3, \tau = 0.25 \) | QR-LASSO | 2.20(3.60) | 44.00 | 4.30 | 0.28 | 0.04 |
| | | QR-E-NET | 2.29(3.68) | 27.00 | 3.33 | 0.14 | 0.04 |
| | | QR-ALASSO | 2.15(3.38) | 60.00 | 4.95 | 0.46 | 0.01 |
| D1: t-distribution | \( \sigma = 1, d = 1, \tau = 0.25 \) | QR-LASSO | 2.32(3.81) | 30.50 | 4.96 | 1.66 | 0.04 |
| | | QR-E-NET | 2.55(3.87) | 29.50 | 4.77 | 1.46 | 0.03 |
| | | QR-ALASSO | 2.36(3.78) | 11.50 | 4.97 | 1.95 | 0.00 |
| | \( \sigma = 3, d = 1, \tau = 0.25 \) | QR-LASSO | 4.70(7.15) | 1.50 | 5.00 | 2.93 | 0.06 |
| | | QR-E-NET | 4.72(7.14) | 1.50 | 4.99 | 2.91 | 0.05 |
| | | QR-ALASSO | 4.68(7.12) | 0.50 | 4.99 | 2.93 | 0.00 |
| | \( \sigma = 1, \tau = 0.50 \) | QR-LASSO | 0.03(2.91) | 40.50 | 4.96 | 1.27 | 0.04 |
| | | QR-E-NET | 0.02(3.27) | 29.00 | 4.64 | 1.10 | 0.03 |
| | | QR-ALASSO | 0.07(3.24) | 32.50 | 4.99 | 1.56 | 0.00 |
| | \( \sigma = 3, \tau = 0.50 \) | QR-LASSO | 0.14(6.89) | 1.50 | 4.99 | 2.85 | 0.05 |
| | | QR-E-NET | 0.14(6.91) | 1.00 | 4.97 | 2.84 | 0.05 |
| | | QR-ALASSO | 0.14(6.84) | 0.50 | 5.00 | 2.86 | 0.00 |
| | \( \sigma = 3, \tau = 0.50 \) | QR-LASSO | 0.13(6.90) | 2.00 | 4.99 | 2.85 | 0.00 |
3.2.2. D2 and D4 under Normal Distribution

Results at the design matrices with collinearity inducing high leverage points (D2 and D4) are shown in Tables 2 and A1. In the unweighted penalized QR case, the adaptive penalized procedures outperform the non-adaptive procedures in all scenarios in all 16 (100%) respective pairwise comparisons with respect to correctly selected (shrink) zero coefficients. However, with respect to the percentage of correctly fitted models, 14 (88%) adaptive penalized procedures outperform their respective non-adaptive procedures and 13 (81%) adaptive penalized procedures outperform their respective non-adaptive procedures with respect to prediction. In the weighted penalized WQR scenarios, WQR-ALASSO and WQR-AE-NET, 10 (63%) adaptive penalized procedures outperform their respective non-adaptive procedures with respect to prediction and 12 (75%) with respect to both the percentage of correctly fitted models and correctly shrunk zero coefficients. The results at $\tau = 0.50$ and at $\tau = 0.25$ RQ levels are similar in D4 (see Table A1 for results at $\tau = 0.50$).

Table 2. Penalized weighted and unweighted quantile regression at D2 and D4 for $n = 50$ ($\tau = 0.25$ and $\tau = 0.50$) under the normal distribution; bold text indicate superior performance.

| Distribution | Parameter | Method  | Median(MAD) | Correctly Fitted | No of Zeros | Median(1) |
|--------------|-----------|---------|-------------|-----------------|-------------|-----------|
| D2 N-distribution | $\sigma = 1, \tau = 0.25$ | QR-LASSO | -1.10(5.65) | 38.00 | 4.00 | 0.01 | 0.03 |
| | | QR-E-NET | -0.93(5.55) | 2.50 | 2.55 | 0.00 | 0.02 |
| | | QR-ALASSO | -1.08(6.00) | 91.50 | 4.92 | 0.01 | 0.02 |
| | | QR-AE-NET | -0.36(5.01) | 81.50 | 4.81 | 0.00 | 0.03 |
| | $\sigma = 3, \tau = 0.25$ | QR-LASSO | 0.96(6.30) | 17.50 | 4.23 | 0.43 | 0.02 |
| | | QR-E-NET | 1.25(8.55) | 9.50 | 3.39 | 0.16 | 0.02 |
| | | QR-ALASSO | 0.90(6.45) | 56.50 | 4.98 | 0.48 | 0.01 |
| | | QR-AE-NET | 1.35(6.64) | 72.50 | 4.93 | 0.22 | 0.01 |
| | $\sigma = 1, \tau = 0.50$ | WQR-LASSO | -1.49(4.49) | 62.50 | 4.47 | 0.00 | 0.03 |
| | | WQR-E-NET | -1.56(6.41) | 11.00 | 3.34 | 0.00 | 0.03 |
| | | WQR-ALASSO | -1.69(4.44) | 94.50 | 4.95 | 0.01 | 0.04 |
| | | WQR-AE-NET | -1.64(6.36) | 92.00 | 4.92 | 0.00 | 0.06 |
| | $\sigma = 3, \tau = 0.25$ | WQR-LASSO | 1.47(7.50) | 57.50 | 4.89 | 0.00 | 0.01 |
| | | WQR-E-NET | 1.50(7.53) | 53.00 | 4.76 | 0.38 | 0.01 |
| | $\sigma = 1, \tau = 0.50$ | QR-LASSO | -1.87(5.78) | 36.00 | 3.83 | 0.01 | 0.02 |
| | | QR-E-NET | -1.91(5.72) | 0.00 | 1.94 | 0.00 | 0.02 |
| | | QR-ALASSO | 0.01(5.76) | 98.50 | 4.99 | 0.00 | 0.03 |
| | | QR-AE-NET | -1.61(5.17) | 94.00 | 4.94 | 0.00 | 0.05 |
| | $\sigma = 3, \tau = 0.50$ | WQR-LASSO | -2.22(4.24) | 66.50 | 4.57 | 0.00 | 0.04 |
| | | WQR-E-NET | -2.04(4.39) | 30.00 | 3.86 | 0.00 | 0.04 |
| | | WQR-ALASSO | -2.29(4.37) | 99.50 | 5.00 | 0.00 | 0.04 |
| | | WQR-AE-NET | -2.20(4.27) | 99.00 | 4.99 | 0.00 | 0.06 |
| | $\sigma = 3, \tau = 0.50$ | WQR-LASSO | -1.06(4.49) | 36.50 | 4.69 | 0.54 | 0.04 |
| | | WQR-E-NET | -0.99(4.28) | 28.00 | 4.34 | 0.37 | 0.05 |
| | | WQR-ALASSO | -0.01(2.39) | 46.50 | 4.85 | 0.48 | 0.01 |
| | | WQR-AE-NET | -0.03(2.49) | 36.00 | 4.61 | 0.39 | 0.01 |
| D4 N-distribution | $\sigma = 1, \tau = 0.25$ | QR-LASSO | 0.97(1.65) | 10.00 | 3.88 | 0.61 | 0.02 |
| | | QR-E-NET | 2.74(3.71) | 1.00 | 1.69 | 0.16 | 0.01 |
| | | QR-ALASSO | 0.88(1.50) | 65.00 | 5.00 | 0.42 | 0.01 |
| | | QR-AE-NET | 2.41(3.44) | 41.00 | 4.54 | 0.16 | 0.01 |
| | $\sigma = 3, \tau = 0.25$ | QR-LASSO | 1.67(6.06) | 3.50 | 4.07 | 1.26 | 0.02 |
| | | QR-E-NET | 2.64(5.30) | 0.00 | 2.25 | 0.75 | 0.01 |
| | | QR-ALASSO | 2.60(7.05) | 0.00 | 4.32 | 1.69 | 0.00 |
| | | QR-AE-NET | 2.83(13) | 0.00 | 2.86 | 1.19 | 0.00 |
| | $\sigma = 1, \tau = 0.25$ | WQR-LASSO | 0.47(0.78) | 64.50 | 4.51 | 0.00 | 0.03 |
| | | WQR-E-NET | 0.49(0.85) | 30.00 | 3.96 | 0.00 | 0.04 |
| | | WQR-ALASSO | 0.52(0.77) | 99.50 | 5.00 | 0.00 | 0.03 |
| | | WQR-AE-NET | 0.56(0.78) | 99.00 | 4.99 | 0.00 | 0.05 |
| | $\sigma = 3, \tau = 0.25$ | WQR-LASSO | 1.37(2.31) | 15.00 | 4.46 | 1.09 | 0.04 |
| | | WQR-E-NET | 1.52(2.38) | 35.00 | 4.29 | 0.28 | 0.04 |
| | | WQR-ALASSO | 1.85(2.63) | 0.00 | 5.00 | 1.58 | 0.01 |
| | | WQR-AE-NET | 1.93(2.72) | 0.00 | 4.96 | 1.51 | 0.02 |
Again, the performance with respect to all measures decreases as the error term distribution becomes heavier, i.e., as $\sigma$ increases, the effectiveness of shrinking zero coefficients correctly tends to be compromised. The pairwise comparisons between the weighted versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) and their respective non-weighted versions (QR-ALASSO and QR-AE-NET) result in the former outperforming the latter in all cases with respect to prediction while doing so in the majority of cases with respect to the other two measures. Overall, the WQR-ALASSO generally outperforms all the other models with respect to all measures.

### 3.2.3. D3 and D5 under the Normal Distribution

Results in Tables 3 and A2 show variable/model selection and prediction performance in the presence of collinearity hiding points (D3 and D5) under the Gaussian distribution. In the unweighted scenarios, the 16 pairwise comparisons demonstrate that the adaptive versions of penalized procedures (QR-ALASSO and QR-AE-NET) dominate their non-adaptive versions as follows; in prediction, they perform better 100% and 88% of the time, respectively; with respect to correctly fitted models they perform better 100% and 50% (while in 50% of the time they perform equally) of the time, respectively; and with respect to correctly shrinking zero coefficients, they both perform better 100% of the time, respectively. The performance picture in the weighted scenario is as follows; with respect to prediction, both WQR-LASSO and WQR-AE-NET outperform their respective adaptive versions 63% of the time; with respect to correctly fitted models, the adaptive versions WQR-ALASSO and WQR-AE-NET outperform their non-adaptive ones 88% of the time, while with respect to correctly shrinking zero coefficients they outperform their non-adaptive ones 100% of the time.

Table 3. Penalized weighted and unweighted quantile regression at D3 and D5 for $n = 50$ ($\tau = 0.25$ and $\tau = 0.50$) under the normal distribution; bold text indicate superior performance.

| Distribution | Parameter | Method          | Median(MAD) Test Error | Correctly Fitted | No of Zeros c.zero inc.zero | Median($\lambda$) |
|--------------|-----------|-----------------|------------------------|-----------------|----------------------------|------------------|
|              | $\sigma = 1, \tau = 0.25$ | QR-LASSO | 0.65(1.21) | 60.50 | 4.45 | 0.00 | 0.01 |
|              |           | QR-E-NET | 0.81(1.46) | 12.50 | 3.48 | 0.00 | 0.01 |
|              |           | QR-ALASSO | 0.79(1.21) | 96.50 | 5.00 | 0.06 | 0.02 |
|              |           | QR-AE-NET | 0.87(1.35) | 99.00 | 5.00 | 0.01 | 0.02 |
|              | $\sigma = 3, \tau = 0.25$ | QR-LASSO | 2.01(3.79) | 42.50 | 4.48 | 0.35 | 0.01 |
|              |           | QR-E-NET | 2.47(4.20) | 17.50 | 3.97 | 0.43 | 0.01 |
|              |           | QR-ALASSO | 2.17(3.49) | 74.50 | 4.75 | 0.19 | 0.00 |
|              |           | QR-AE-NET | 2.74(4.16) | 44.50 | 4.48 | 0.42 | 0.00 |
|              | $\sigma = 1, \tau = 0.25$ | WQR-LASSO | 0.48(0.91) | 28.50 | 3.63 | 0.00 | 0.04 |
|              |           | WQR-E-NET | 0.41(0.77) | 15.50 | 3.59 | 0.00 | 0.03 |
|              |           | WQR-ALASSO | 0.48(0.70) | 98.00 | 4.99 | 0.01 | 0.04 |
|              |           | WQR-AE-NET | 0.50(0.82) | 98.50 | 4.99 | 0.01 | 0.04 |
|              | $\sigma = 3, \tau = 0.25$ | WQR-LASSO | 1.25(2.12) | 29.00 | 4.39 | 0.57 | 0.03 |
|              |           | WQR-E-NET | 1.46(2.31) | 24.50 | 4.21 | 0.42 | 0.04 |
|              |           | WQR-ALASSO | 0.24(4.81) | 40.00 | 4.90 | 0.77 | 0.02 |
|              |           | WQR-AE-NET | 0.67(4.38) | 44.00 | 4.86 | 0.68 | 0.04 |
|              | $\sigma = 1, \tau = 0.50$ | QR-LASSO | −0.01(1.19) | 62.50 | 4.48 | 0.00 | 0.02 |
|              |           | QR-E-NET | −0.04(1.39) | 12.50 | 3.45 | 0.00 | 0.02 |
|              |           | QR-ALASSO | 0.02(1.15) | 99.50 | 5.00 | 0.00 | 0.02 |
|              |           | QR-AE-NET | −0.01(1.43) | 97.50 | 4.98 | 0.00 | 0.03 |
|              | $\sigma = 3, \tau = 0.50$ | QR-LASSO | −0.08(3.61) | 53.00 | 4.54 | 0.30 | 0.02 |
|              |           | QR-E-NET | −0.10(4.02) | 11.00 | 3.92 | 0.54 | 0.01 |
|              |           | QR-ALASSO | 0.06(3.55) | 74.00 | 4.80 | 0.10 | 0.00 |
|              |           | QR-AE-NET | 0.05(3.96) | 18.50 | 4.17 | 0.35 | 0.00 |
Table 3. Cont.

| Distribution | Parameter | Method       | Median(MAD) | Correctly Fitted | No of Zeros |
|--------------|-----------|--------------|-------------|-----------------|-------------|
|              |           | Test Error   |              | c.zero           | inc.zero    | Median(λ) |
| $\sigma = 1, \tau = 0.50$ | WQR-LASSO | $-0.01(0.67)$ | 71.50       | 4.65            | 0.00        | 0.04       |
|              | WQR-E-NET | $-0.01(0.77)$ | 35.00       | 3.96            | 0.00        | 0.05       |
|              | WQR-ALASSO| 0.01(0.78)    | 100.00      | 5.00            | 0.00        | 0.04       |
|              | WQR-AE-NET| 0.01(0.81)    | 100.00      | 5.00            | 0.00        | 0.07       |
| $\sigma = 3, \tau = 0.50$ | WQR-LASSO | $-0.05(2.27)$ | 39.50       | 4.66            | 0.48        | 0.04       |
|              | WQR-E-NET | $-0.05(2.30)$ | 31.50       | 4.38            | 0.35        | 0.04       |
|              | WQR-ALASSO| $1.04(4.52)$  | 51.00       | 4.96            | 0.72        | 0.03       |
|              | WQR-AE-NET| $0.94(4.15)$  | 63.00       | 4.98            | 0.48        | 0.04       |
| $\sigma = 1, \tau = 0.50$ | QR-LASSO  | $-0.06(2.34)$ | 54.50       | 4.98            | 0.80        | 0.00       |
|              | QR-E-NET  | $-0.02(4.67)$ | 0.00        | 4.97            | 2.91        | 0.09       |
|              | QR-ALASSO | 0.01(1.46)    | 95.50       | 5.00            | 0.07        | 0.00       |
|              | QR-AE-NET | 0.00(4.66)    | 0.00        | 4.99            | 2.87        | 0.09       |
| $\sigma = 3, \tau = 0.50$ | QR-LASSO  | 0.08(5.40)    | 11.50       | 4.98            | 2.56        | 0.01       |
|              | QR-E-NET  | $-0.03(5.42)$ | 0.00        | 4.96            | 2.92        | 0.09       |
|              | QR-ALASSO | 0.07(4.74)    | 26.50       | 5.00            | 1.70        | 0.00       |
|              | QR-AE-NET | $-0.02(5.47)$ | 0.00        | 4.99            | 2.95        | 0.01       |
| $\sigma = 1, \tau = 0.50$ | WQR-LASSO | 0.010(0.70)   | 68.50       | 4.61            | 0.00        | 0.04       |
|              | WQR-E-NET | 0.010(0.69)   | 34.50       | 3.95            | 0.00        | 0.04       |
|              | WQR-ALASSO| 0.01(0.72)    | 100.00      | 5.00            | 0.00        | 0.04       |
|              | WQR-AE-NET| 0.000(0.73)   | 98.00       | 4.98            | 0.00        | 0.06       |
| $\sigma = 3, \tau = 0.50$ | WQR-LASSO | $0.04(2.26)$  | 42.00       | 4.72            | 0.47        | 0.04       |
|              | WQR-E-NET | $0.03(2.26)$  | 29.00       | 4.35            | 0.38        | 0.05       |
|              | WQR-ALASSO| 0.00(2.28)    | 17.50       | 4.99            | 0.91        | 0.01       |
|              | WQR-AE-NET| $-0.01(2.39)$ | 13.00       | 5.00            | 0.93        | 0.02       |

Thus, the adaptive weights improve the performance of the models with respect to all measures in the unweighted scenario while doing so with respect to correctly fitted models and correctly shrinking zero coefficients in the unweighted scenario. However, in the weighted scenario, the adaptive weights hamper the performance of the models with respect to prediction. Although the $\tau = 0.25$ at D5 scenario is not included in the tables, we obtain similar results as at $\tau = 0.50$ (see Table A2).

3.2.4. D2 and D3 under the $t$-Distribution

The heavy-tailed $t$-distribution scenario results for D2 and D3 with collinearity hiding and inducing observations, respectively, are shown in Tables 4, 5, A3 and A4; respectively. For brevity, the results of the heavy-tailed D4 and D5 scenarios are not reported since they are similar to those of D2 and D3.

At the predictor matrix with collinearity-inducing points (D2), under the $t$-distribution, all adaptive versions outperformed their non-adaptive versions with respect to all metrics. In the unweighted scenarios, the 16 pairwise comparisons clearly demonstrate that the adaptive versions of penalized procedures ($QR$-ALASSO and $QR$-AE-NET) dominate their non-adaptive versions as follows; in prediction, they each perform better 88% of the time over their respective non-adaptive versions; with respect to correctly fitted models, they perform better 100% of the time and 38% (while in 62% of the time they perform equally) over their respective non-adaptive versions; and with respect to correctly shrinking zero coefficients, they outperform their non-adaptive ones 100% and 88% of the time over their respective non-adaptive versions.

In the weighted scenarios, the 16 pairwise comparisons clearly demonstrate that the adaptive versions of penalized procedures ($WQR$-ALASSO and $WQR$-AE-NET) dominate their non-adaptive versions as follows; in prediction, they each perform better 100% and 75% of the time over their respective non-adaptive versions; with respect to correctly fitted
models, they perform better 88% of the time each, respectively; and with respect to correctly shrinking zero coefficients, they outperform their non-adaptive ones 88% and 100% of the time over their respective non-adaptive versions.

Table 4. Penalized weighted and unweighted quantile regression at $D_2$ for $n = 50$ ($\tau = 0.25$ and $\tau = 0.50$) under the heavy tailed $t$-distribution.

| Distribution | Parameter | Method | Median(MAD) | Correctly Fitted | No of Zeros | Median(\lambda) |
|--------------|-----------|--------|-------------|------------------|-------------|-----------------|
| $\sigma = 1, d = 1, \tau = 0.25$ | | QR-LASSO | 2.49(4.51) | 1.50 | 4.26 | 1.50 | 0.02 |
| | | QR-E-NET | 3.26(5.05) | 0.00 | 2.79 | 1.32 | 0.02 |
| | | QR-AE-LASSO | 2.43(3.87) | 24.00 | 4.93 | 1.52 | 0.01 |
| | | QR-AE-NET | 2.71(4.29) | 0.00 | 3.03 | 0.78 | 0.01 |
| $\sigma = 0.50, d = 1, \tau = 0.25$ | | QR-LASSO | 1.08(2.02) | 15.50 | 4.05 | 0.54 | 0.02 |
| | | QR-E-NET | 1.75(3.09) | 0.00 | 1.30 | 0.20 | 0.02 |
| | | QR-AE-LASSO | 1.16(2.16) | 52.00 | 4.88 | 0.76 | 0.02 |
| | | QR-AE-NET | 1.45(2.38) | 0.50 | 3.27 | 0.32 | 0.03 |
| $\sigma = 1, d = 1, \tau = 0.25$ | | WQR-LASSO | 1.57(2.51) | 21.50 | 4.85 | 1.65 | 0.04 |
| | | WQR-E-NET | 1.61(2.60) | 4.00 | 4.35 | 1.55 | 0.04 |
| | | WQR-ALASSO | 1.49(2.63) | 8.00 | 4.79 | 2.12 | 0.00 |
| | | WQR-AE-NET | 1.60(2.71) | 1.00 | 4.57 | 2.16 | 0.00 |
| $\sigma = 0.50, d = 1, \tau = 0.25$ | | WQR-LASSO | 0.81(1.41) | 46.50 | 4.74 | 0.82 | 0.04 |
| | | WQR-E-NET | 0.89(1.56) | 5.00 | 3.77 | 0.74 | 0.04 |
| | | WQR-ALASSO | 0.70(1.33) | 55.00 | 5.00 | 0.95 | 0.02 |
| | | WQR-AE-NET | 0.79(1.42) | 51.00 | 4.90 | 0.88 | 0.03 |
| $\sigma = 1, d = 1, \tau = 0.50$ | | QR-LASSO | 0.00(3.64) | 3.00 | 3.97 | 1.20 | 0.02 |
| | | QR-E-NET | −0.01(4.32) | 0.00 | 2.67 | 0.92 | 0.02 |
| | | QR-ALASSO | 0.08(3.26) | 42.50 | 4.97 | 1.16 | 0.01 |
| | | QR-AE-NET | 0.02(3.85) | 0.50 | 3.34 | 0.74 | 0.01 |
| $\sigma = 0.50, d = 1, \tau = 0.50$ | | QR-LASSO | 0.02(1.70) | 30.00 | 4.37 | 0.53 | 0.03 |
| | | QR-E-NET | 0.01(2.49) | 0.00 | 1.64 | 0.27 | 0.02 |
| | | QR-ALASSO | 0.04(1.64) | 77.00 | 5.00 | 0.54 | 0.02 |
| | | QR-AE-NET | 0.01(1.98) | 2.00 | 3.34 | 0.33 | 0.02 |
| $\sigma = 1, d = 1, \tau = 0.50$ | | WQR-LASSO | 0.03(2.21) | 29.50 | 4.73 | 1.14 | 0.04 |
| | | WQR-E-NET | 0.02(2.30) | 9.00 | 4.11 | 1.00 | 0.04 |
| | | WQR-ALASSO | 0.02(2.21) | 30.50 | 4.94 | 1.52 | 0.01 |
| | | WQR-AE-NET | 0.05(2.30) | 17.00 | 4.58 | 1.40 | 0.01 |
| $\sigma = 0.50, d = 1, \tau = 0.50$ | | WQR-LASSO | 0.01(1.15) | 51.50 | 4.67 | 0.51 | 0.04 |
| | | WQR-E-NET | 0.01(1.25) | 7.00 | 3.51 | 0.46 | 0.04 |
| | | WQR-ALASSO | 0.00(1.11) | 67.00 | 5.00 | 0.66 | 0.04 |
| | | WQR-AE-NET | 0.00(1.22) | 69.50 | 4.98 | 0.59 | 0.07 |
| $\sigma = 1, d = 6, \tau = 0.25$ | | QR-LASSO | 0.82(1.30) | 14.00 | 3.40 | 0.01 | 0.02 |
| | | QR-E-NET | 1.18(1.80) | 0.00 | 1.90 | 0.00 | 0.02 |
| | | QR-ALASSO | 0.84(1.30) | 90.00 | 4.90 | 0.01 | 0.02 |
| | | QR-AE-NET | 1.04(1.53) | 0.00 | 2.91 | 0.00 | 0.02 |
| $\sigma = 0.50, d = 6, \tau = 0.25$ | | QR-LASSO | 0.40(0.64) | 16.00 | 3.51 | 0.01 | 0.02 |
| | | QR-E-NET | 0.62(1.02) | 0.00 | 1.90 | 0.00 | 0.02 |
| | | QR-ALASSO | 0.38(0.63) | 98.00 | 4.98 | 0.00 | 0.02 |
| | | QR-AE-NET | 0.63(1.03) | 0.00 | 2.28 | 0.00 | 0.02 |
| $\sigma = 1, d = 6, \tau = 0.25$ | | WQR-LASSO | 0.50(0.97) | 42.00 | 4.10 | 0.00 | 0.04 |
| | | WQR-E-NET | 0.51(1.00) | 1.50 | 2.72 | 0.00 | 0.04 |
| | | WQR-ALASSO | 0.55(0.91) | 92.00 | 4.92 | 0.01 | 0.03 |
| | | WQR-AE-NET | 0.60(0.97) | 56.50 | 4.45 | 0.00 | 0.04 |
| $\sigma = 0.50, d = 6, \tau = 0.25$ | | WQR-LASSO | 0.26(0.49) | 31.00 | 3.92 | 0.00 | 0.04 |
| | | WQR-E-NET | 0.27(0.51) | 2.50 | 2.62 | 0.00 | 0.04 |
| | | WQR-ALASSO | 0.28(0.45) | 98.50 | 4.99 | 0.00 | 0.04 |
| | | WQR-AE-NET | 0.30(0.49) | 83.50 | 4.82 | 0.00 | 0.05 |
The pairwise comparisons between the weighted versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) and their unweighted versions (QR-ALASSO and QR-AE-NET) result in the former outperforming the latter in all respective cases with respect to prediction, while there is no clear “winner” with respect to the other two measures. Overall, the WQR-ALASSO generally outperforms all other models with respect to all measures.

At the predictor matrix with collinearity-masking points (D3) under the t-distribution, all adaptive versions outperformed their non-adaptive versions with respect to all metrics. In the unweighted scenarios, the 16 pairwise comparisons clearly demonstrate that the adaptive versions of penalized procedures (QR-ALASSO and QR-AE-NET) dominate their non-adaptive versions as follows; in prediction, they perform better 63% and 75% of the time over their respective non-adaptive versions; with respect to correctly fitted models, they perform better 100% of the time and 38% (while in 62% of the time they perform equally) over their respective non-adaptive versions; and with respect to correctly shrinking zero coefficients, they outperform their non-adaptive ones 100% of the time in both cases over their respective non-adaptive versions.

In the weighted scenarios, the 16 pairwise comparisons of the adaptive versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) against their non-adaptive versions show that the dominance of the former class of models is somewhat reduced, though they generally outperform the latter. In the prediction scenario, the adaptive versions are outperformed by their non-adaptive versions 62% of the time in both respective cases, while it is the opposite with respect to correctly fitted models and with respect to correctly shrinking zero coefficients. With respect to these two metrics, the adaptive versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) dominate their non-adaptive versions as follows; with respect to correctly fitted models, they perform better 62% of the time, while with respect to correctly shrinking zero coefficients they outperform their non-adaptive ones 100% of the time in both cases over their respective non-adaptive versions.

The pairwise comparisons between the weighted versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) and their unweighted versions (QR-ALASSO and QR-AE-NET) result in the former outperforming the latter in all respective cases with respect to prediction while doing so marginally better with respect to the other two measures. Overall, the WQR-ALASSO generally outperforms all other models with respect to all measures.

3.2.5. D6 under the t-Distribution

The results of variable/model selection as well as prediction performance at the design matrix with collinearity and high leverage points (D6) under heavy-tailed distributions (t-distributions with $d \in (1; 6)$ degrees of freedom) scenarios are shown in Tables 6 and A5. In the unweighted scenarios, the 16 pairwise comparisons clearly demonstrate that the adaptive versions of penalized procedures (QR-ALASSO and QR-AE-NET) dominate their non-adaptive versions as follows; in prediction, they each perform better 75% and 62% of the time over their respective non-adaptive versions; with respect to correctly fitted models, they perform better in 100% and 88% of the time over their respective non-adaptive versions, while with respect to correctly shrinking zero coefficients they outperform their non-adaptive ones 100% of the time in both cases over their respective non-adaptive versions.

In the weighted scenarios, the 16 pairwise comparisons clearly demonstrate that the adaptive versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) dominate their non-adaptive versions as follows; with respect to prediction, they outperform their non-adaptive ones 100% of the time in both cases over their respective non-adaptive versions; and with respect to both correctly fitted models and correctly shrinking zero coefficients they both perform better in 88% of the time over their respective non-adaptive versions.
Table 5. Penalized weighted and unweighted quantile regression at $D_3$ for $n = 50$ ($\tau = 0.25$ and $\tau = 0.50$) under the heavy tailed $t$-distribution; bold text indicate superior performance.

| Distribution | Parameter | Method | Median(MAD) | Test Error | Correctly Fitted | No of Zeros | Median (λ) |
|--------------|-----------|--------|-------------|-------------|------------------|-------------|------------|
| $\sigma = 1, d = 1, \tau = 0.25$ | | QR-LASSO | 1.32(2.39) | 17.00 | 3.84 | 0.43 | 0.03 |
| | | QR-E-NET | 2.07(3.49) | 0.00 | 1.77 | 0.30 | 0.02 |
| | | QR-ALASSO | 1.84(3.21) | 44.50 | 4.91 | 0.74 | 0.01 |
| | | QR-AE-NET | 2.23(3.39) | 2.50 | 3.69 | 0.29 | 0.02 |
| $\sigma = 0.50, d = 1, \tau = 0.25$ | | QR-LASSO | 0.77(1.28) | 56.00 | 4.55 | 0.25 | 0.03 |
| | | QR-E-NET | 1.75(2.65) | 0.00 | 2.17 | 0.20 | 0.02 |
| | | QR-ALASSO | 0.90(1.56) | 84.50 | 4.99 | 0.27 | 0.03 |
| | | QR-AE-NET | 1.44(2.04) | 2.00 | 4.01 | 0.18 | 0.04 |
| $\sigma = 1, d = 1, \tau = 0.50$ | | WQR-LASSO | 0.67(1.30) | 67.50 | 4.55 | 0.54 | 0.04 |
| | | WQR-E-NET | 0.77(1.47) | 17.00 | 4.45 | 0.47 | 0.04 |
| | | WQR-ALASSO | 0.90(1.56) | 21.50 | 4.99 | 0.27 | 0.03 |
| | | WQR-AE-NET | 0.91(1.56) | 22.50 | 4.93 | 0.77 | 0.04 |
| $\sigma = 1, d = 6, \tau = 0.50$ | | WQR-LASSO | −0.02(2.39) | 28.50 | 4.30 | 0.46 | 0.06 |
| | | WQR-E-NET | −0.07(3.56) | 10.00 | 4.45 | 0.25 | 0.02 |
| | | WQR-ALASSO | 0.02(3.05) | 27.50 | 4.81 | 0.86 | 0.01 |
| | | WQR-AE-NET | 0.04(3.43) | 0.00 | 3.08 | 0.48 | 0.02 |
| $\sigma = 0.50, d = 1, \tau = 0.25$ | | WQR-LASSO | 0.01(1.23) | 63.00 | 4.71 | 0.31 | 0.03 |
| | | WQR-E-NET | 0.00(2.15) | 0.00 | 2.08 | 0.21 | 0.02 |
| | | WQR-ALASSO | 0.01(1.47) | 79.00 | 5.00 | 0.37 | 0.04 |
| | | WQR-AE-NET | 0.04(2.33) | 0.00 | 3.08 | 0.48 | 0.02 |
| $\sigma = 0.50, d = 1, \tau = 0.50$ | | WQR-LASSO | 0.01(1.23) | 63.00 | 4.71 | 0.31 | 0.03 |
| | | WQR-E-NET | 0.00(2.15) | 0.00 | 2.08 | 0.21 | 0.02 |
| | | WQR-ALASSO | 0.01(1.47) | 79.00 | 5.00 | 0.37 | 0.04 |
| | | WQR-AE-NET | 0.04(2.33) | 0.00 | 3.08 | 0.48 | 0.02 |
| $\sigma = 0.50, d = 6, \tau = 0.25$ | | WQR-LASSO | −0.01(1.17) | 69.50 | 4.88 | 0.49 | 0.04 |
| | | WQR-E-NET | −0.01(1.29) | 10.50 | 3.36 | 0.40 | 0.04 |
| | | WQR-ALASSO | 0.01(0.42) | 98.50 | 4.99 | 0.61 | 0.03 |
| | | WQR-AE-NET | 0.02(1.26) | 73.50 | 4.98 | 0.50 | 0.05 |
| $\sigma = 0.50, d = 6, \tau = 0.50$ | | WQR-LASSO | 0.00(0.33) | 47.50 | 4.26 | 0.00 | 0.05 |
| | | WQR-E-NET | 0.00(0.36) | 4.00 | 2.84 | 0.00 | 0.04 |
| | | WQR-ALASSO | 0.01(0.42) | 95.50 | 4.99 | 0.00 | 0.05 |
| | | WQR-AE-NET | 0.01(0.43) | 93.00 | 4.93 | 0.00 | 0.07 |
Table 6. Penalized weighted and unweighted quantile regression at D6 under the heavy-tailed t-distribution for n = 50 (τ = 0.25 and τ = 0.50); bold text indicate superior performance.

| Distribution Parameter | Method            | Median(MAD) Test Error | Correctly Fitted | No of Zeros c.zero inc.zero | Mediant(λ) |
|------------------------|-------------------|------------------------|------------------|----------------------------|------------|
| σ = 1, d = 1, τ = 0.25 | QR-LASSO          | 2.54(4.43)             | 22.50            | 4.71                       | 1.34       | 0.04         |
|                        | QR-E-NET          | 2.79(4.69)             | 3.00             | 3.06                       | 0.96       | 0.04         |
|                        | QR-ALASSO         | 2.43(4.26)             | 19.50            | 4.74                       | 1.19       | 0.02         |
|                        | QR-AE-NET         | 2.71(4.66)             | 0.00             | 4.07                       | 1.07       | 0.02         |
| σ = 0.50, d = 1, τ = 0.25 | QR-LASSO       | 1.10(1.95)             | 25.50            | 4.15                       | 1.02       | 0.03         |
|                        | QR-E-NET          | 1.34(2.37)             | 0.00             | 2.36                       | 0.44       | 0.03         |
|                        | QR-ALASSO         | 1.15(2.14)             | 37.00            | 4.73                       | 0.77       | 0.03         |
|                        | QR-AE-NET         | 1.27(2.39)             | 8.00             | 4.22                       | 0.45       | 0.04         |
| σ = 1, d = 1, τ = 0.25 | WQR-LASSO         | 1.63(3.05)             | 14.00            | 4.66                       | 1.52       | 0.04         |
|                        | WQR-E-NET         | 1.82(3.22)             | 1.00             | 3.61                       | 1.07       | 0.04         |
|                        | WQR-ALASSO        | 1.60(2.86)             | 14.00            | 4.94                       | 1.95       | 0.02         |
|                        | WQR-AE-NET        | 1.75(3.06)             | 11.50            | 4.80                       | 1.79       | 0.04         |
| σ = 0.50, d = 1, τ = 0.25 | WQR-LASSO       | 1.06(1.99)             | 67.00            | 4.80                       | 0.38       | 0.04         |
|                        | WQR-E-NET         | 1.27(2.23)             | 1.00             | 2.45                       | 0.32       | 0.03         |
|                        | WQR-ALASSO        | 0.80(1.43)             | 10.00            | 3.95                       | 0.69       | 0.00         |
|                        | WQR-AE-NET        | 0.96(1.71)             | 0.50             | 2.79                       | 0.66       | 0.00         |
| σ = 1, d = τ = 0.50   | QR-LASSO          | −0.06(3.48)            | 37.00            | 4.69                       | 0.90       | 0.04         |
|                        | QR-E-NET          | −0.06(3.72)            | 1.50             | 2.67                       | 0.76       | 0.04         |
|                        | QR-ALASSO         | −0.02(3.40)            | 41.50            | 4.91                       | 0.87       | 0.02         |
|                        | QR-AE-NET         | 0.06(3.82)             | 2.00             | 3.99                       | 0.58       | 0.03         |
| σ = 0.50, d = 1, τ = 0.50 | QR-LASSO        | −0.04(4.12)            | 33.00            | 4.30                       | 0.36       | 0.04         |
|                        | QR-E-NET          | 0.00(1.83)             | 0.00             | 2.45                       | 0.33       | 0.03         |
|                        | QR-ALASSO         | −0.01(1.81)            | 64.50            | 4.96                       | 0.63       | 0.04         |
|                        | QR-AE-NET         | 0.01(1.99)             | 16.00            | 4.23                       | 0.31       | 0.06         |
| σ = 1, d = 1, τ = 0.50 | WQR-LASSO        | 0.02(1.52)             | 58.50            | 4.67                       | 0.33       | 0.04         |
|                        | WQR-E-NET         | 0.03(1.76)             | 0.10             | 2.92                       | 0.29       | 0.04         |
|                        | WQR-ALASSO        | −0.01(1.19)            | 69.00            | 4.86                       | 0.43       | 0.00         |
|                        | WQR-AE-NET        | 0.00(1.41)             | 7.50             | 3.21                       | 0.40       | 0.00         |
| σ = 0.50, d = 1, τ = 0.50 | WQR-LASSO        | 0.82(1.25)             | 53.50            | 4.26                       | 0.00       | 0.04         |
|                        | WQR-E-NET         | 0.86(1.35)             | 3.00             | 2.44                       | 0.00       | 0.00         |
|                        | WQR-ALASSO        | 0.81(1.23)             | 96.00            | 4.96                       | 0.00       | 0.03         |
|                        | WQR-AE-NET        | 0.88(1.29)             | 38.50            | 4.22                       | 0.00       | 0.04         |
| σ = 1, d = 6, τ = 0.25 | QR-LASSO          | 0.40(0.66)             | 63.50            | 4.51                       | 0.00       | 0.00         |
|                        | QR-E-NET          | 0.41(0.72)             | 2.50             | 2.77                       | 0.00       | 0.04         |
|                        | QR-ALASSO         | 0.38(0.62)             | 97.50            | 4.97                       | 0.00       | 0.03         |
|                        | QR-AE-NET         | 0.40(0.66)             | 26.00            | 4.15                       | 0.00       | 0.05         |
| σ = 0.50, d = 6, τ = 0.25 | WQR-LASSO       | 0.54(0.92)             | 43.00            | 4.23                       | 0.00       | 0.04         |
|                        | WQR-E-NET         | 0.55(0.96)             | 2.00             | 3.07                       | 0.00       | 0.04         |
|                        | WQR-ALASSO        | 0.37(0.69)             | 98.00            | 4.98                       | 0.00       | 0.03         |
|                        | WQR-AE-NET        | 0.40(0.73)             | 59.00            | 4.52                       | 0.00       | 0.05         |
| σ = 1, d = 6, τ = 0.25 | WQR-LASSO        | 0.26(0.47)             | 35.50            | 4.06                       | 0.00       | 0.04         |
|                        | WQR-E-NET         | 0.25(0.51)             | 5.00             | 3.16                       | 0.00       | 0.03         |
|                        | WQR-ALASSO        | 0.25(0.42)             | 99.50            | 5.00                       | 0.00       | 0.04         |
|                        | WQR-AE-NET        | 0.27(0.43)             | 83.50            | 4.84                       | 0.00       | 0.06         |

The pairwise comparisons between the weighted versions of penalized procedures (WQR-ALASSO and WQR-AE-NET) and their non-weighted versions (QR-ALASSO and QR-AE-NET) result in former outperforming the latter in all cases with respect to prediction while there is no clear “winner” with respect to the other two measures. Overall, the WQR-ALASSO generally outperforms all other models with respect to all measures.

3.3. Examples

In this section, we illustrate the efficacy of the weighted QR-RIDGE-based adaptive weights in penalized QR using two data sets from the literature often used to illustrate the efficacy of robust methodologies in mitigation against collinearity influential as well as high leverage points in general, viz., the Jet-Turbine Engine [32] and Gunst and Mason [33] data sets. The Jet-Turbine Engine data set, which has high collinearity reducing points, is very popular with Engineers (see [32,34]). In contrast, the Gunst and Mason data set, which has collinearity-inducing points, is very popular with statisticians (see also [33]).
3.4. The Jet-Turbine Engine Data

The Jet-Turbine Engine data set [32] data set consists of 40 observations with thrust of a jet-turbine engine as the response variable ($Y$) and predictor variables: primary speed of rotation ($X_1$), secondary speed of rotation ($X_2$), fuel flow rate ($X_3$), pressure ($X_4$), exhaust temperature ($X_5$), and ambient temperature at time of test ($X_6$). According to [34], observations 6 and 20 are high leverage collinearity reducing points. The red data is standardized to correlation form and the predictor variable is generated by $Y_1 = X_1' \beta_1 + \epsilon_1$, $\epsilon_1 \sim t_6$ for the 38 observations excluding the collinearity reducing ones with $X_2$ and $Y_2$ comprising cases (observations) 6 and 20, such that $X = (X_1, X_2)'$ and $Y = (Y_1, Y_2)'$, where $\beta_1 = (50, 0, 0, 10, 15, 0)'$.

Table 7 summarizes the estimated $QR$ $\beta$s and their biases from the true coefficients/true $\beta$s ($\beta_1 = (50, 0, 0, 10, 15, 0)'$). In the unweighted scenario, the zero coefficients are shrunk to zero 33% of the time with no clear “winner” between the adaptive and non-adaptive penalization. In the weighted scenario, the proportion of correctly shrunk coefficients increases to 54%, again with no clear “winner” between the adaptive and non-adaptive penalization. However, the adaptive penalization seem to guard against incorrectly shrinking non-zero coefficients to zero (see $\beta_4$ under $WQR-LASSO$ at $\tau = 0.50$). As expected, the $LASSO$ (and $ALASSO$) penalties outperform $E-NET$ (and $AE-NET$) penalties. From Figure 2a, it is clear that the weighting increases the variability of the residuals as more outliers are flagged by the weighting.

![Figure 2](image_url)

**Figure 2.** Residual box plots for data sets at $\tau = 0.50$ $QR$ level. (a) Box plots for the Gunst and Mason data set and (b) box plots for the Jet-Turbine Engine data set.

3.5. The Gunst and Mason Data

The performance of the proposed adaptive procedures is assessed in this Section using the [33] data set. The data set has 49 observations (countries) with the response variable gross national product ($GNP$) and predictor variables: infant death rate ($TNFD$), physician to population ratio ($PHYS$), population density ($DENS$), density of agricultural land ($AGDS$), measure of literacy ($LIT$), and higher education index ($HIED$) corresponding the response variable $Y$ and predictor variables and predictor variables $X_1$, $X_2$, $X_3$, $X_4$, $X_5$, and $X_6$, respectively. From the literature, strong collinearity between the $DENS$ ($X_3$) and $AGDS$ ($X_4$) exist, and a few observations are high leverage points. The following points are considered outlying and influential: 7 (Canada), 13 (El Salvador), 17 (Hong Kong), 20 (India), 39 (Singapore), and 46 (United States of America) with observations 17 and 39 being high leverage collinearity inducing points. We standardize the data to correlation form and generate the predictor variable as $Y_1 = X_1' \beta_1 + \epsilon_1$, $\epsilon_1 \sim t_6$ for the first 43 observations (without influential observations 7, 13, 17, 20, 39 and 46) with $X_2$ and $Y_2$.
comprising observations 7, 13, 17, 20, 39 and 46, such that $X = (X_1, X_2)'$ and $Y = (Y_1, Y_2)'$, where $\beta_1 = (0, 8, -13, 0, 0, 6)'$.

Table 7. Penalized QR parameter estimation results for the Jet-Turbine Engine data set [32]; bold text indicate superior performance.

|     | $\beta$ | NON-BIASED $\beta$ (Bias) | QR-LASSO $\beta$ (Bias) | QR-E-NET $\beta$ (Bias) | QR-ALASSO $\beta$ (Bias) | QR-AE-NET $\beta$ (Bias) |
|-----|---------|---------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| $\tau = 0.25$ | $\lambda$ | -0.72 | -1.11(0.39) | -0.60(-0.12) | -0.69(-0.03) | -0.72(0.00) | -0.67(-0.05) |
|     | intercept | -0.01 | 0.01 | 0.07 | 0.02 |
|     | $X_1$ | 50.00 | 15.15(-34.85) | 33.20(-16.80) | 21.68(-28.37) | 17.55(-32.45) | 24.53(-25.47) |
|     | $X_2$ | 0.00 | 84.78(64.78) | 10.34(10.34) | 14.14(14.14) | 26.40(26.40) | 27.17(27.17) |
|     | $X_3$ | 0.00 | -89.20(-89.20) | 0.00(0.00) | 0.00(0.00) | -20.53(-20.53) | -26.10(-26.10) |
|     | $X_4$ | 10.00 | 28.94(18.94) | 14.18(14.18) | 19.49(9.49) | 27.98(17.98) | 25.27(15.27) |
|     | $X_5$ | 15.00 | 30.54(15.54) | 16.51(1.51) | 18.57(3.57) | 22.19(7.19) | 22.54(7.54) |
|     | $X_6$ | 0.00 | 10.21(10.21) | -4.28(-4.28) | -4.75(-4.75) | -0.76(-0.76) | 0.57(0.57) |
| $\tau = 0.50$ | $\lambda$ | 0.00 | 0.08 | 0.03 | 0.32 | 0.10 |
|     | intercept | 0.00 | 0.40(0.40) | 0.44(0.44) | 0.39(0.39) | 0.57(0.57) | 0.53(0.53) |
|     | $X_1$ | 50.00 | 38.47(-11.53) | 9.68(-40.32) | 16.28(-33.72) | 27.98(-22.02) | 16.09(-33.91) |
|     | $X_2$ | 0.00 | 41.62(41.62) | 33.30(33.30) | 17.72(17.72) | 35.97(35.97) | 21.51(21.51) |
|     | $X_3$ | 0.00 | -37.77(-37.77) | 0.00(0.00) | 8.14(8.14) | 0.00(0.00) | 0.00(0.00) |
|     | $X_4$ | 10.00 | 12.73(2.73) | 21.57(11.57) | 18.00(8.00) | 5.43(-4.57) | 21.60(11.60) |
|     | $X_5$ | 15.00 | 19.06(4.06) | 6.02(-8.98) | 12.42(-2.58) | 1.56(-13.44) | 13.52(-1.48) |
|     | $X_6$ | 0.00 | 4.21(4.21) | 0.00(0.00) | -3.06(-3.06) | 0.00(0.00) | -0.90(-0.90) |

Note that these models with no constant term, the intercept $\beta_0 = 0.0$ and $-0.72$ among the adaptive penalization procedures. Thus, adaptive penalization clearly outperforms non-adaptive penalization. In the weighted scenario, the proportion of correctly shrunk coefficients increases to 83% with 67% shrunk to zero among the adaptive penalization procedures and 50% shrunk to zero among the non-adaptive penalization procedures. Therefore, the adaptive penalization clearly outperforms non-adaptive penalization. In the weighted scenario, the proportion of correctly shrunk coefficients increases to 83% with 67% shrunk to zero among the adaptive penalization procedures. Thus, adaptive penalization clearly outperforms non-adaptive penalization. Additionally, in the weighted scenario the ALASSO and AE-NET perform equally. As expected the LASSO (and ALASSO) penalties outperform E-NET (and AE-NET) penalties. From Figure 2b, it is clear that the weighting increases the variability of the residuals as more outliers are flagged by the weighting.

Table 8 summarizes the estimated QR $\beta$s and their biases from the true coefficients/true $\beta$s ($\beta_1 = (0, 8, -13, 0, 0, 6)'$). In the unweighted scenario, the zero coefficients are shrunk to zero 38% of the time with 25% shrunk to zero among the adaptive penalization procedures and 100% shrunk to zero among the non-adaptive penalization procedures. Thus, adaptive penalization clearly outperforms non-adaptive penalization. In the weighted scenario, the proportion of correctly shrunk coefficients increases to 83% with 67% shrunk to zero among the adaptive penalization procedures and 100% shrunk to zero among the non-adaptive penalization procedures. Thus, adaptive penalization clearly outperforms non-adaptive penalization. Additionally, in the weighted scenario the ALASSO and AE-NET perform equally. As expected the LASSO (and ALASSO) penalties outperform E-NET (and AE-NET) penalties. From Figure 2b, it is clear that the weighting increases the variability of the residuals as more outliers are flagged by the weighting.
Table 8. Penalized QR parameter estimation results for the Gunst and Mason data set; bold text indicate superior performance.

|       | $\beta$  | $\beta_{(Bias)}$ | $\beta_{(Bias)}$ | $\beta_{(Bias)}$ | $\beta_{(Bias)}$ | $\beta_{(Bias)}$ |
|-------|-----------|------------------|------------------|------------------|------------------|------------------|
|       |           | NON-BIASED       | QR-LASSO         | QR-E-NET         | QR-ALASSO        | QR-AE-NET        |
| $\lambda$ |          |                  |                  |                  |                  |                  |
| intercept | $-0.72$  | $0.38(-1.10)$    | $-0.85(0.13)$    | $-0.85(0.13)$    | $-0.98(0.26)$    | $-0.69(0.03)$    |
| $X_1$  | $0.00$    | $-0.14(0.14)$    | $-0.02(0.02)$    | $0.00(0.00)$     | $0.00(0.00)$     | $0.00(0.00)$     |
| $X_2$  | $8.00$    | $9.29(-1.29)$    | $4.37(0.63)$     | $4.89(3.11)$     | $2.97(5.03)$     | $7.73(0.27)$     |
| $X_3$  | $-13.00$  | $-10.97(-2.03)$  | $0.00(-13.00)$   | $-2.00(-11.00)$  | $-6.35(-6.65)$   | $-9.59(-3.41)$   |
| $X_4$  | $0.00$    | $11.54(-11.54)$  | $0.87(-0.87)$    | $3.01(-3.01)$    | $6.68(-6.68)$    | $10.09(-10.09)$  |
| $X_5$  | $0.00$    | $6.21(-6.21)$    | $1.62(-1.62)$    | $2.09(-2.09)$    | $0.00(0.00)$     | $4.90(-4.90)$    |
| $X_6$  | $6.00$    | $1.61(4.39)$     | $2.23(2.78)$     | $2.19(3.81)$     | $3.30(2.70)$     | $1.74(4.26)$     |
| $\tau = 0.25$ |          |                  |                  |                  |                  |                  |
| $\lambda$ | $0.01$    | $0.02$           | $0.01$           | $0.01$           | $0.01$           | $0.01$           |
| intercept | $0.00$    | $0.44(-0.44)$    | $0.03(-0.03)$    | $0.03(-0.03)$    | $0.14(-0.14)$    | $0.07(-0.07)$    |
| $X_1$  | $0.00$    | $0.72(-0.72)$    | $0.40(-0.40)$    | $0.41(-0.41)$    | $0.00(0.00)$     | $0.00(0.00)$     |
| $X_2$  | $8.00$    | $9.87(-1.87)$    | $10.13(-2.13)$   | $10.13(-2.13)$   | $7.83(0.17)$     | $8.21(-0.21)$    |
| $X_3$  | $-13.00$  | $-5.15(-7.85)$   | $-4.80(-8.20)$   | $-4.80(-8.20)$   | $-0.99(-12.02)$  | $-8.41(-4.59)$   |
| $X_4$  | $0.00$    | $4.68(-4.68)$    | $4.38(-4.38)$    | $4.38(-4.38)$    | $0.00(0.00)$     | $6.89(-6.89)$    |
| $X_5$  | $0.00$    | $3.64(-3.64)$    | $3.71(-3.71)$    | $3.71(-3.71)$    | $0.00(0.00)$     | $0.81(-0.81)$    |
| $X_6$  | $6.00$    | $4.94(1.06)$     | $4.92(1.08)$     | $4.91(1.09)$     | $6.24(-0.24)$    | $4.31(1.69)$     |
| $\tau = 0.50$ |          |                  |                  |                  |                  |                  |
| $\lambda$ | $0.04$    | $0.05$           | $0.05$           | $0.06$           |                  |                  |
| intercept | $0.00$    | $0.44(-0.44)$    | $-0.01(0.01)$    | $0.00(0.00)$     | $0.00(0.00)$     | $-0.01(0.01)$    |
| $X_1$  | $0.00$    | $0.72(-0.72)$    | $0.00(0.00)$     | $0.00(0.00)$     | $0.00(0.00)$     | $0.00(0.00)$     |
| $X_2$  | $8.00$    | $9.87(-1.87)$    | $7.48(0.52)$     | $6.27(1.73)$     | $8.37(-0.37)$    | $8.44(-0.44)$    |
| $X_3$  | $-13.00$  | $-5.15(-7.85)$   | $-5.75(-7.25)$   | $-5.99(-7.01)$   | $-9.29(-3.71)$   | $-9.67(-3.33)$   |
| $X_4$  | $0.00$    | $4.68(-4.68)$    | $0.00(0.00)$     | $0.00(0.00)$     | $0.00(0.00)$     | $0.00(0.00)$     |
| $X_5$  | $0.00$    | $3.64(-3.64)$    | $-0.18(-0.18)$   | $-1.19(1.19)$    | $0.00(0.00)$     | $0.00(0.00)$     |
| $X_6$  | $6.00$    | $4.94(1.06)$     | $10.68(-4.68)$   | $11.21(-5.21)$   | $11.38(-5.38)$   | $11.32(-5.32)$   |

Note that these models with no constant term, the intercept $F^{-1} + \beta_0$ translate to 0 and $-0.72$ under the $t_6$ error term distribution, at quantile levels $\tau = 0.50$ and $\tau = 0.25$, respectively.

4. Discussion

In this article, we proposed ALASSO and AE-NET penalized QR variable selection and parameter estimation procedures (both weighted and unweighted). We used the proposed adaptive weights, $\omega_j$, based on the $\tilde{F}(\tau)$, a WQR-RIDGE parameter estimate in our proposed adaptive penalized QR procedures. Our proposed QR-ALASSO, QR-AE-NET, WQR-ALASSO and WQR-AE-NET variable selection and parameter estimation procedures were subjected to a simulation study. We first discuss the results of this simulation study results and the applications to well-known data sets from the literature.

The simulation study results under the Gaussian distribution at D1-D6 (predictor matrices with collinearity influentiel points, collinearity and high leverage points) show that QR-ALASSO, QR-AE-NET, WQR-ALASSO and WQR-AE-NET outperform their non-adaptive counterparts in at least 63% of the time in each respective scenario when low standard errors of the error term are present as in [30]. This is expected as the adaptive penalized QR methods satisfy sparsity and are asymptotically normally distributed, and provide a balance between bias-variance trade-offs. However, the performance with respect to all measures becomes compromised as the error term distribution becomes heavier. This same pattern of results is also evident in the heavy-tailed $t$-distribution scenario. In unweighted scenario, the QR-ALASSO is superior to other penalized procedures in the presence of larger errors in agreeing with the case of [16]’s results, where ALASSO is superior to LASSO. A similar pattern is observed in the weighted scenarios. However,
a comparison between the weighted scenarios and the unweighted ones shows that the MCD weights enhance the efficacy of the penalized procedures at predictor matrices with collinearity influential observations, collinearity and high leverage points (as demonstrated in [14]) with the WQR-ALASSO performing best.

The applications of the suggested procedures to real life data sets are more or less in line with the simulation studies with the procedures exhibiting a higher efficacy at the predictor matrix with collinearity inducing points, i.e., the [33] data set, than the one with collinearity reducing points, i.e., the [32] data set. However, it must be noted that the latter data set has the severest collinearity inherent in it. The introduction of collinearity reducing to “clean” data only managed to lower the condition number from 52.09 to 47.78, which is indicative of the fact that the data is still highly collinear (see [34]). Again, the WQR-ALASSO performs best as in the simulation studies. Thus, although variable selection and regularization is adversely affected by heavy-tailed error term distributions (outliers), the adaptive weights generally improve the efficacy of penalized QR procedures with the MCD-based weights tending to increase their efficacy. We recommend the use of our ALASSO penalty for variable selection, with an option to use the AE-NET penalty if the objective of the study is prediction. Parameter estimation results from popular data sets in the literature support the applicability of our proposed methods.

**Author Contributions:** Conceptualization, I.M. and E.R.; methodology, I.M.; validation, I.M. and E.R.; formal analysis, I.M.; investigation, I.M.; resources, I.M.; data curation, I.M.; writing—original draft preparation, I.M.; writing—review and editing, E.R.; visualization, E.R.; supervision, E.R.; project administration, I.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by the University of South Africa.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable

**Acknowledgments:** We acknowledge support from University of South Africa and Vaal University of Technology.

**Conflicts of Interest:** The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Abbreviations**

The following abbreviations are used in this manuscript:

| Abbreviation | Description |
|--------------|-------------|
| LS           | least squares |
| LAD          | least absolute deviation |
| QR           | quantile regression |
| LASSO        | least absolute shrinkage and selection operator |
| E-NET        | elastic net |
| SCAD         | smoothly clipped absolute deviation |
| MCP          | minimax concave penalty |
| WQR          | weighted quantile regression |
| ARIDGE       | adaptive RIDGE |
| QR-ALASSO    | quantile regression adaptive LASSO |
| QR-AE-NET    | quantile regression adaptive elastic net |
| WQR-RIDGE    | weighted quantile regression RIDGE |
| WQR-ALASSO   | weighted quantile regression adaptive LASSO |
| WQR-AE-NET   | weighted quantile regression adaptive elastic net |
| MCD          | minimum covariance determinant |
| WRR          | weighted RIDGE regression |
| RQ           | regression quantile |
Appendix A

Table A1. Penalized weighted and unweighted quantile regression at $D_4$ for $n = 50$ ($\tau = 0.50$) under the normal distribution.

| Distribution | Parameter | Method          | Median(MAD) Test Error | Correctly Fitted | No of Zeros | c.zero | inc.zero | Median($\lambda$) |
|--------------|-----------|-----------------|------------------------|-------------------|-------------|--------|---------|------------------|
| $\sigma = 1, \tau = 0.50$ | QR-LASSO   | $-0.03(1.40)$  | 5.50                   | 3.49              | 0.44        | 0.02   |
|               | QR-E-NET  | $-0.07(3.56)$  | 0.00                   | 1.06              | 0.01        | 0.02   |
|               | QR-ALASSO  | $0.00(1.37)$   | 88.50                  | 5.00              | 0.15        | 0.01   |
|               | QR-AE-NET  | $-0.03(3.45)$  | 55.00                  | 4.00              | 0.02        | 0.02   |
| $\sigma = 3, \tau = 0.50$ | QR-LASSO   | $-1.31(6.57)$  | 5.00                   | 3.46              | 0.64        | 0.02   |
|               | QR-E-NET  | $-0.78(5.30)$  | 0.00                   | 1.60              | 0.36        | 0.02   |
|               | QR-ALASSO  | $-1.21(7.80)$  | 0.00                   | 4.97              | 1.62        | 0.00   |
|               | QR-AE-NET  | $-0.27(5.07)$  | 0.00                   | 4.31              | 1.50        | 0.01   |

Table A2. Penalized weighted and unweighted quantile regression at $D_5$ for $n = 50$ ($\tau = 0.25$) under the normal distribution.

| Distribution | Parameter | Method          | Median(MAD) Test Error | Correctly Fitted | No of Zeros | c.zero | inc.zero | Median($\lambda$) |
|--------------|-----------|-----------------|------------------------|-------------------|-------------|--------|---------|------------------|
| $\sigma = 1, \tau = 0.25$ | QR-LASSO   | 2.68(4.14)      | 12.00                  | 5.00              | 1.89        | 0.01   |
|               | QR-E-NET  | 3.27(4.67)      | 0.00                   | 4.94              | 2.94        | 0.04   |
|               | QR-ALASSO  | 1.13(1.85)      | 74.00                  | 5.00              | 0.30        | 0.01   |
|               | QR-AE-NET  | 3.26(4.65)      | 0.00                   | 5.00              | 2.87        | 0.04   |
| $\sigma = 3, \tau = 0.25$ | QR-LASSO   | 3.61(5.43)      | 0.00                   | 5.00              | 2.57        | 0.05   |
|               | QR-E-NET  | 3.68(5.45)      | 0.00                   | 4.94              | 2.87        | 0.04   |
|               | QR-ALASSO  | 3.42(5.20)      | 12.50                  | 5.00              | 2.03        | 0.00   |
|               | QR-AE-NET  | 3.54(5.47)      | 0.00                   | 4.98              | 2.94        | 0.00   |
| $\sigma = 1, \tau = 0.25$ | WQR-LASSO  | 0.50(0.82)      | 74.00                  | 5.00              | 1.89        | 0.01   |
|               | WQR-E-NET  | 0.53(0.82)      | 46.00                  | 4.23              | 2.94        | 0.04   |
|               | WQR-ALASSO | 0.51(0.79)      | 99.50                  | 5.00              | 0.01        | 0.04   |
|               | WQR-AE-NET | 0.52(0.79)      | 99.50                  | 5.00              | 0.01        | 0.04   |
| $\sigma = 3, \tau = 0.25$ | WQR-LASSO  | 1.57(2.39)      | 36.50                  | 4.64              | 0.55        | 0.04   |
|               | WQR-E-NET  | 1.67(2.46)      | 27.50                  | 4.24              | 0.43        | 0.04   |
|               | WQR-ALASSO | 1.59(2.36)      | 41.50                  | 4.96              | 0.67        | 0.02   |
|               | WQR-AE-NET | 1.71(2.42)      | 50.50                  | 4.97              | 0.54        | 0.03   |
Table A3. Penalized weighted and unweighted quantile regression at $D_2$ for $n = 50$ ($\tau = 0.50$) under the heavy tailed t-distribution.

| Distribution | Parameter | Method | Median (MAD) | Correctly Fitted | No of Zeros | Median ($\lambda$) |
|--------------|-----------|--------|--------------|------------------|-------------|------------------|
| $\sigma = 1, d = 6, \tau = 0.50$ | QR-LASSO | $-0.01(1.27)$ | 11.50 | 3.31 | 0.00 | 0.02 |
| | QR-E-NET | $-0.02(1.60)$ | 0.00 | 1.89 | 0.00 | 0.02 |
| | QR-ALASSO | $0.031(2.23)$ | 87.00 | 4.87 | 0.00 | 0.02 |
| | QR-AE-NET | 0.051(1.43) | 0.00 | 3.05 | 0.00 | 0.03 |
| $\sigma = 0.50, d = 6, \tau = 0.50$ | QR-LASSO | $-0.02(0.61)$ | 10.50 | 3.32 | 0.00 | 0.03 |
| | QR-E-NET | 0.00(0.84) | 0.00 | 1.94 | 0.00 | 0.02 |
| | QR-ALASSO | $-0.02(0.61)$ | 97.00 | 4.97 | 0.00 | 0.02 |
| | QR-AE-NET | 0.01(0.77) | 0.00 | 3.00 | 0.00 | 0.03 |

Table A4. Penalized weighted and unweighted quantile regression at $D_3$ for $n = 50$ ($\tau = 0.25$) under the heavy tailed t-distribution.

| Distribution | Parameter | Method | Median (MAD) | Correctly Fitted | No of Zeros | Median ($\lambda$) |
|--------------|-----------|--------|--------------|------------------|-------------|------------------|
| $\sigma = 1, d = 6, \tau = 0.25$ | QR-LASSO | $0.82(1.29)$ | 3.50 | 2.91 | 0.02 | 0.01 |
| | QR-E-NET | 1.57(2.23) | 0.00 | 0.17 | 0.00 | 0.01 |
| | QR-ALASSO | $0.811(2.43)$ | 96.50 | 4.97 | 0.00 | 0.03 |
| | QR-AE-NET | 0.01(1.23) | 97.00 | 4.97 | 0.00 | 0.03 |
| $\sigma = 0.50, d = 6, \tau = 0.25$ | WQR-LASSO | $0.37(0.66)$ | 4.00 | 2.84 | 0.00 | 0.01 |
| | WQR-E-NET | 0.40(1.66) | 0.00 | 0.14 | 0.00 | 0.01 |
| | WQR-ALASSO | $0.400(0.64)$ | 100.00 | 5.00 | 0.00 | 0.03 |
| | WQR-AE-NET | 0.79(1.09) | 0.00 | 3.00 | 0.00 | 0.05 |

Table A5. Penalized weighted and unweighted quantile regression at $D_6$ under the heavy tailed t-distribution for $n = 50$ ($\tau = 0.50$).

| Distribution | Parameter | Method | Median (MAD) | Correctly Fitted | No of Zeros | Median ($\lambda$) |
|--------------|-----------|--------|--------------|------------------|-------------|------------------|
| $\sigma = 1, d = 6, \tau = 0.50$ | QR-LASSO | 0.05(1.26) | 51.00 | 4.35 | 0.00 | 0.05 |
| | QR-E-NET | 0.03(1.30) | 1.50 | 2.29 | 0.00 | 0.04 |
| | QR-ALASSO | $0.031(2.23)$ | 99.50 | 5.00 | 0.00 | 0.04 |
| | QR-AE-NET | 0.051(1.28) | 51.50 | 4.51 | 0.00 | 0.07 |
| $\sigma = 0.50, d = 6, \tau = 0.50$ | QR-LASSO | 0.00(0.64) | 72.50 | 4.65 | 0.00 | 0.05 |
| | QR-E-NET | 0.00(1.07) | 5.00 | 3.00 | 0.00 | 0.04 |
| | QR-ALASSO | 0.000(6.63) | 96.50 | 4.96 | 0.00 | 0.04 |
| | QR-AE-NET | 0.000(6.64) | 24.50 | 4.00 | 0.00 | 0.06 |
| $\sigma = 1, d = 6, \tau = 0.50$ | WQR-LASSO | 0.010(0.67) | 55.50 | 4.45 | 0.00 | 0.04 |
| | WQR-E-NET | $-0.01(0.89)$ | 6.00 | 3.23 | 0.00 | 0.04 |
| | WQR-ALASSO | $-0.010(6.33)$ | 94.50 | 4.93 | 0.00 | 0.04 |
| | WQR-AE-NET | $-0.010(0.66)$ | 23.00 | 3.91 | 0.00 | 0.05 |
| $\sigma = 0.50, d = 6, \tau = 0.50$ | WQR-LASSO | $-0.01(0.44)$ | 44.00 | 4.19 | 0.00 | 0.04 |
| | WQR-E-NET | 0.000(4.79) | 5.00 | 3.13 | 0.00 | 0.04 |
| | WQR-ALASSO | 0.000(3.79) | 99.00 | 4.99 | 0.00 | 0.05 |
| | WQR-AE-NET | 0.000(3.79) | 85.00 | 4.85 | 0.00 | 0.07 |
References

1. Chatterjee, S.; Hadi, A.S. Impact of simultaneous omission of a variable and an observation on a linear regression equation. *Comput. Stat. Data Anal.* **1988**, *6*, 129–144. [CrossRef]

2. Bloomfield, P.; Steiger, W. Least absolute deviations curve-fitting. *SIAM J. Sci. Stat. Comput.* **1980**, *1*, 290–301. [CrossRef]

3. Koenker, R.; Bassett, G., Jr. Regression quantiles. *Econom. J. Econom. Soc.* **1978**, pp. 33–50. [CrossRef]

4. Breiman, L. Better subset regression using the nonnegative garrote. *Technometrics* **1995**, *37*, 373–384. [CrossRef]

5. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]

6. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **1996**, *58*, 267–288. [CrossRef]

7. Zou, H.; Hastie, T. Regularization and variable selection via the Elastic Net. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 301–320. [CrossRef]

8. Yuan, M.; Lin, Y. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B* **2006**, *68*, 49–67. [CrossRef]

9. Zhao, P.; Yu, B. On model selection consistency of Lasso. *J. Mach. Learn. Res.* **2006**, *7*, 2541–2563.

10. Fan, J.; Li, R. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *J. Am. Stat. Assoc.* **2001**, *96*, 1348–1360. [CrossRef]

11. Zhang, C.H. Nearly unbiased variable selection under minimax concave penalty. *Ann. Stat.* **2010**, *38*, 894–942. [CrossRef]

12. Karunanuni, R.J.; Kong, L.; Tu, W. Efficient robust doubly adaptive regularized regression with applications. *Stat. Methods Med. Res.* **2019**, *28*, 2210–2226. [CrossRef] [PubMed]

13. Salibián-Barrera, M.; Wei, Y. Weighted quantile regression with nonelliptically structured covariates. *Can. J. Stat.* **2008**, *36*, 595–611. [CrossRef]

14. Ranganai, E.; Mudjombo, I. Variable Selection and Regularization in Quantile Regression via Minimum Covariance Determinant Based Weights. *Entropy* **2021**, *23*, 33. [CrossRef] [PubMed]

15. Fan, J.; Peng, H. Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* **2004**, *32*, 928–961. [CrossRef]

16. Zou, H. The Adaptive Lasso and Its Oracle Properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429. [CrossRef]

17. Frommlet, F.; Nuel, G. An adaptive ridge procedure for l0 regularization. *PLoS ONE* **2016**, *11*, e0148620. [CrossRef]

18. Wu, Y.; Liu, Y. Variable selection in quantile regression. *Stat. Sin.* **2009**, *19*, 801–817.

19. Zou, H.; Zhang, H.H. On the adaptive elastic-net with a diverging number of parameters. *Ann. Stat.* **2009**, *37*, 1733. [CrossRef]

20. Rousseauw, P. Multivariate Estimation with High Breakdown Point. *Math. Stat. Appl.* **1985**, *8*, 283–297. [CrossRef]

21. Zhao, Q. Restricted regression quantiles. *J. Multivar. Anal.* **2000**, *72*, 78–99. [CrossRef]

22. Ranganai, E. Aspects of Model Development Using Regression Quantiles and Elemental Regressions. Ph.D. Thesis, Stellenbosch University, South Africa, 2007.

23. Gilioni, A.; Simonoff, J.S.; Sengupta, B. Robust weighted LAD regression. *Comput. Stat. Data Anal.* **2006**, *50*, 3124–3140. [CrossRef]

24. Hubert, M.; Rousseauw, P.J. Robust regression with both continuous and binary regressors. *J. Stat. Plan. Inference* **1997**, *57*, 153–163. [CrossRef]

25. Pollard, D. Asymptotics for Least Absolute Deviation Regression Estimators. *Econom. Theory* **1991**, *7*, 186–199. [CrossRef]

26. Knight, K. Asymptotics for L1-estimators of regression parameters under heteroscedasticityY. *Can. J. Stat.* **1999**, *27*, 497–507. [CrossRef]

27. Koenker, R. *Quantile Regression*; Econometric Society Monographs, Cambridge University Press: Cambridge, UK, 2005.

28. Jongh, P.D.; de Wet, T.; Welsh, A.H. Mallows-Type Bounded-Influence-Regression Trimmed Means. *J. Am. Stat. Assoc.* **1988**, *83*, 805–810. [CrossRef]

29. Ranganai, E.; Vuuren, J.O.V.; Wet, T.D. Multiple Case High Leverage Diagnosis in Regression Quantiles. *J. Am. Stat. Assoc.* **1988**, *83*, 805–810. [CrossRef]

30. Arslan, O. Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression. *Comput. Stat. Data Anal.* **2012**, *56*, 1952–1965. [CrossRef]

31. Yi, C. R-Add-On-Software: Hqreg. Available online: https://cloud.r-project.org/web/packages/hqreg (accessed on 20 February 2022).

32. Brownlee, K.A. Statistical theory and methodology in science and engineering. In *A Wiley Publication in Applied Statistics*; Wiley: Hoboken, NJ, USA, 1965.

33. Ruppert, D.; Carroll, R.J. Trimmed least squares estimation in the linear model. *J. Am. Stat. Assoc.* **1980**, *75*, 828–838. [CrossRef]

34. Cook, R.D. Influential Observations in Linear Regression. *J. Am. Stat. Assoc.* **1979**, *74*, 169–174. [CrossRef]