

11-1-2013

Robust Regression Estimators When There are Tied Values

Rand R. Wilcox

University of Southern California, Los Angeles, rwilcox@usc.edu

Florence Clark

University of Southern California, Los Angeles, fclark@osot.usc.edu

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Wilcox, Rand R. and Clark, Florence (2013) "Robust Regression Estimators When There are Tied Values," *Journal of Modern Applied Statistical Methods*: Vol. 12 : Iss. 2 , Article 3.

DOI: 10.22237/jmasm/1383278520

Invited Article: **Robust Regression Estimators When There are Tied Values**

Rand R. Wilcox

University of Southern California
Los Angeles, CA

Florence Clark

University of Southern California
Los Angeles, CA

It is well known that when using the ordinary least squares regression estimator, outliers among the dependent variable can result in relatively poor power. Many robust regression estimators have been derived that address this problem, but the bulk of the results assume that the dependent variable is continuous. It is demonstrated that when there are tied values, several robust regression estimators can perform poorly in terms of controlling the Type I error probability, even with a large sample size. The presence of tied values does not necessarily mean that they perform poorly, but there is the issue of whether there is a robust estimator that performs reasonably well in situations where other estimators do not. The main result is that a modification of the Theil–Sen estimator achieves this goal. Results on the small-sample efficiency of the modified Theil–Sen estimator are reported as well. Data from the Well Elderly 2 Study, which motivated this study, are used to illustrate that the modified Theil–Sen estimator can make a practical difference.

Keywords: Tied values, Harrell–Davis estimator, MM–estimator, Coakley–Hettmansperger estimator, rank-based regression, Theil–Sen estimator, Well Elderly II Study, perceived control

Introduction

It is well known that the ordinary least squares (OLS) regression estimator is not robust (e.g., Hampel et al., 1987; Huber & Ronchetti, 2009; Maronna et al. 2006; Staudte & Sheather, 1990; Wilcox, 2012a, b). One concern is that even a single outlier among the values associated with the dependent variable can result in relatively poor power. Numerous robust regression estimators have been derived that are aimed at dealing with this issue, a fairly comprehensive list of which can be found in Wilcox (2012b, Chapter 10). But the bulk of the published

Rand R. Wilcox is a Professor of Psychology. Email him at: rwilcox@usc.edu. Florence Clark is a Professor of Occupational Science and Occupational Therapy. Email her at: fclark@osot.usc.edu.

results on robust regression estimators assume the dependent variable is continuous.

Motivated by data stemming from the Well II study (Jackson et al. 2009), this paper examines the impact of tied values on the probability of a Type I error when testing hypotheses via various robust regression estimators. Many of the dependent variables in the Well Elderly study were the sum of Likert scales. Consequently, with a sample size of 460, tied values were inevitable. Moreover, the dependent variables were found to have outliers, suggesting that power might be better using a robust estimator. But given the goal of testing the hypothesis of a zero slope, it was unclear whether the presence of tied values might impact power and the probability of a Type I error.

Preliminary simulations indicated that indeed there is a practical concern. Consider, for example, the Theil (1950) and Sen (1968) estimator. One of the dependent variables (CESD) in the Well Elderly study reflected a measure of depressive symptoms. It consists of the sum of twenty Likert scales with possible scores ranging between 0 and 60. The actual range of scores in the study was 0 to 56. Using the so-called MAD-median rule (e.g., Wilcox, 2012b), 5.9% of the values were flagged as outliers, raising concerns about power despite the relatively large sample size. A simulation was run where observations were randomly sampled with replacement from the CESD scores and the independent variable was taken to be values randomly sampled from a standard normal distribution and independent of the CESD scores. The estimated Type I error probability, when testing at the .05 level, was .002 based on 2000 replications. A similar result was obtained when the dependent variable was a measure of perceived control. Now 7.8% of the values are declared outliers. As an additional check, the values for the dependent variable were generated from a beta-binomial distribution having probability function

$$P(Y = y) = \frac{B(m - y + r, y + s)}{(m + 1)B(m - y + 1, y + 1)B(r, s)}, \quad (1)$$

where B is the complete beta function and the sample space consists of the integers $0, \dots, m$. For $r = s = 1$ as well as $(r, s) = (1, 9)$, again, the actual level was less than .01.

Other robust estimators were found to have a similar problem or situations were encountered where they could not be computed. The estimators that were considered included Yohai's (1987) MM-estimator, the one-step estimator derived

by Agostinelli and Markatou (1998), Rousseeuw's (1984) least trimmed squares (LTS) estimator, the Coakley and Hettmansperger (1993) M-estimator, the Koenker and Bassett (1978) quantile estimator and a rank-based estimator stemming from Jaeckel (1972). The MM-estimator and the LTS estimator were applied via the R package `robustbase`, the Agostinelli—Markatou estimator was applied with the R package `wle`, the quantile regression estimator was applied via the R package `quantreg`, the rank-based estimator was applied using the R package `Rfit`, and the Coakley—Hettmansperger and Theil—Sen estimators were applied via the R package `WRS`. A percentile bootstrap method was used to test the hypothesis of a zero slope, which allows heteroscedasticity and has been found to perform relatively well, in terms of controlling the probability of a Type I error, compared to other strategies that have been studied (Wilcox, 2012b). The MM-estimator, the Agostinelli—Markatou estimator and the Coakley—Hettmansperger estimator routinely terminated in certain situations due to some computational issue. This is not to suggest that they always performed poorly, this is not the case. But when dealing a skewed discrete distribution (a beta-binomial distribution with $m = 10$, $r = 9$ and $s = 1$), typically a p-value could not be computed. The other estimators had estimated Type I errors well below the nominal level. The R package `Rfit` includes a non-bootstrap test of the hypothesis that the slope is zero. Again the actual level was found to be substantially less than the nominal level in various situations, and increasing n only made matters worse. So this raised the issue of whether any reasonably robust estimator can be found that avoids the problems just described.

For completeness, when dealing with discrete distributions, an alternative approach is to use multinomial logistic regression. This addresses an issue that is potentially interesting and useful. But in the Well study, for example, what was deemed more relevant was modeling the typical CESD score given a value for CAR. That is, a regression estimator that focuses on some conditional measure of location, given a value for the independent variable, was needed.

The goal in this paper is to suggest a simple modification of the Theil—Sen estimator that avoids the problems just indicated. Section 2 reviews the Theil—Sen estimator and indicates why it can be highly unsatisfactory. Then the proposed modification is described. Section 3 describes the hypothesis testing method that is used. Section 4 summarizes simulation estimates of the actual Type I error probability when testing at the .05 level and it reports some results on its small-sample efficiency. Section 5 uses data from Well Elderly II study to illustrate that the modified Theil—Sen estimator can make a substantial practical difference.

The Theil–Sen Estimator and the Suggested Modification

When the dependent variable is continuous, the Theil–Sen estimator enjoys good theoretical properties and it performs well in simulations in terms of power and Type I error probabilities when testing hypotheses about the slope (e.g., Wilcox, 2012b). Its mean squared error and small-sample efficiency compare well to the OLS estimator as well as other robust estimators that have been derived (Dietz, 1987; Wilcox, 1998). Dietz (1989) established that its asymptotic breakdown point is approximately .29. Roughly, about 29% of the points must be changed in order to make the estimate of the slope arbitrarily large or small. Other asymptotic properties have been studied by Wang (2005) and Peng et al. (2008). Akritas et al. (1995) applied it to astronomical data and Fernandes and Leblanc (2005) to remote sensing. Although the bulk of the results on the Theil–Sen estimator deal with situations where the dependent variable is continuous, an exception is the paper by Peng et al. (2008) that includes results when dealing a discontinuous error term. They show that when the distribution of the error term is discontinuous, the Theil–Sen estimator can be super-efficient. They establish that even in the continuous case, the slope estimator may or may not be asymptotically normal. Peng et al. also establish the strong consistency and the asymptotic distribution of the Theil–Sen estimator for a general error distribution. Currently, a basic percentile bootstrap seems best when testing hypotheses about the slope and intercept, which has been found to perform well even when the error term is heteroscedastic (e.g., Wilcox, 2012b).

The Theil–Sen estimate of the slope is the usual sample median based on all of the slopes associated with any two distinct points. Consequently, practical concerns previously outlined are not surprising in light of results when dealing with inferential methods based on the sample median (Wilcox, 2012a, section 4.10.4). Roughly, when there are tied values, the sample median is not asymptotically normal. Rather, as sample size increases, the cardinality of its sample can decrease, which in turn creates concerns about the more obvious methods for testing hypotheses

Recent results on comparing quantiles (Wilcox et al., 2013) suggest a modification that might deal the concerns previously indicated: replace the usual sample median with the Harrell and Davis (1982) estimate of the median, which uses a weighted average of all the order statistics.

To describe the computational details, let $(Y_1, X_1), \dots, (Y_n, X_n)$ be a random sample from some unknown bivariate distribution. Assuming that $X_j \neq X_k$ for any $j < k$, let

ROBUST REGRESSION ESTIMATORS WHEN THERE ARE TIED VALUES

$$b_{jk} = \frac{Y_j - Y_k}{X_j - X_k}, 1 \leq j < k \leq n.$$

The Theil–Sen estimate of the slope, $\hat{\beta}_1$, is taken to be the usual sample median based on the b_{jk} values. The intercept is typically estimated with $\hat{\beta}_0 = M_y - \hat{\beta}_1 M_x$, where M_y is the usual sample median based on Y_1, \dots, Y_n . This will be called the TS estimator henceforth.

For notational convenience, Let Z_1, \dots, Z_ℓ denote the b_{jk} values, where $\ell = (n^2 - n) / 2$. Let U be a random variable having a beta distribution with parameters $a = (\ell + 1)q$ and $b = (\ell + 1)(1 - q)$, $0 < q < 1$. Let

$$W_i = P\left(\frac{i-1}{\ell} \leq U \leq \frac{i}{\ell}\right).$$

Let $Z_{(1)} \leq \dots \leq Z_{(\ell)}$ denote the Z_1, \dots, Z_ℓ values written in ascending order. The Harrell and Davis (1982) estimate of the q th quantile is

$$\hat{\theta}_q = \sum W_i Z_{(i)}$$

Consequently, estimate the slope with $\tilde{\beta}_1 = \hat{\theta}_q$. The intercept is estimated with the Harrell–Davis estimate of the median based on $Y_1 - \tilde{\beta}_1 X_1, \dots, Y_n - \tilde{\beta}_1 X_n$. This will be called the HD estimator.

So the strategy is to avoid the problem associated with the usual sample median by using a quantile estimator that results in a sampling distribution that in general does not have tied values. Because the Harrell–Davis estimator uses all of the order statistics, the expectation is that in general it accomplishes this goal. For the situations described in the introduction, for example, no tied values were found among the 5000 estimates of the slope. This, in turn, offers some hope that good control over the probability of a Type I error can be achieved via a percentile bootstrap method.

It is noted that alternative quantile estimators have been proposed that are also based on a weighted average of all the order statistics. In terms of its standard error, Sfakianakis and Verginis (2006) show that in some situations the Harrell–Davis estimator competes well with alternative estimators that again use a

weighted average of all the order statistics, but there are exceptions. Additional comparisons of various estimators are reported by Parrish (1990), Sheather and Marron (1990), as well as Dielman, Lowry and Pfaffenberger (1994). Perhaps one of these alternative estimators offers some practical advantage for the situation at hand, but this is not pursued here.

Hypothesis Testing

As previously indicated, a percentile bootstrap method has been found to be an effective way of testing hypotheses based on a robust regression estimators, including situations where the error term is heteroscedastic (e.g., Wilcox, 2012b). Also, because it is unclear when the HD estimator is asymptotically normal, using a percentile bootstrap method for the situation at hand seems preferable compared to using some pivotal test statistic based on some estimate of the standard error. (For general theoretical results on the percentile bootstrap method that are relevant here, see Liu & Singh, 1997.)

When testing

$$H_0 : \beta_1 = 0, \tag{2}$$

the percentile bootstrap begins by resampling with replacement n vectors of observations from $(Y_1, X_1), \dots, (Y_n, X_n)$ yielding say $(Y_1^*, X_1^*), \dots, (Y_n^*, X_n^*)$. Based on this bootstrap sample, let $\tilde{\beta}_1^*$ be the resulting estimate of the slope. Repeat this process B times yielding $\tilde{\beta}_{1b}^*, 1, \dots, B$. Let A be the proportion of $\tilde{\beta}_{1b}^*$ values that are less than null value, 0, and let C be the number of times $\tilde{\beta}_{1b}^*$ is equal to the null value. Then a (generalized) p-value when testing (2) is

$$p = 2 \min(\hat{p}, 1 - \hat{p}),$$

where $\hat{p} = \frac{A}{B} + .5 \frac{C}{B}$. Here, $B = 599$ is used. This choice appears to work well with robust estimators in terms of controlling the probability of a Type I error (e.g., Wilcox, 2012b). However, based on results in Racine and MacKinnon (2007), $B > 599$ might provide improved power.

Simulation Results

Simulations were used to study the small-sample properties of the HD estimator. When comparing the small-sample efficiency of estimators, 4000 replications were used with $n = 20$. When estimating the actual probability of a Type I error, 2000 replications were used with sample sizes 20 and 60. Some additional simulations were run with $n = 200$ as a partial check on the R functions that were used to apply the methods.

To ensure tied values, values for Y were generated from one of four discrete distributions. The first two were beta-binomial distributions. Here $m = 10$ is used in which case the possible values for Y are the integers 0, 1, ..., 10. The idea is to consider a situation where the number of tied values is relatively large. The values for r and s were taken to be $(r,s) = (1,9)$, which is a skewed distribution with mean 1, and $r = s = 3$, which is a symmetric distribution with mean 5. The third distribution was a discretized version of the normal distribution. More precisely, n observations were generated from a standard normal distribution, say V_1, \dots, V_n , and Y_i is taken to be $2V_i$ rounded to the nearest integer. (Among the 4,000 replications, the observed values for Y ranged between -9 and 10.) This process for generating observations will be labeled SN. For the final distribution, observations were generated as done in SN but with a standard normal replace by a contaminated normal having distribution

$$H(y) = .9\Phi(y) + .1\Phi\left(\frac{y}{10}\right),$$

where $\Phi(y)$ is a standard normal distribution. The contaminated normal has mean zero and variance 10.9. It is heavy-tailed, roughly meaning that it tends to generate more outliers than the normal distribution. This process will be labeled CN.

Estimated Type I error probabilities are shown in Table 1 for $n = 20$ and 60 when testing at the $\alpha = .05$ level. In Table 1, $B(r,s,m)$ indicates that Y has a beta-binomial distribution. The column headed by TS shows the results when using the Theil–Sen estimator. Notice that the estimates are substantially less than the nominal level when $n = 20$. Moreover, the estimated level actually decreases when n is increased to 60. In contrast, when using the HD estimator, the estimated level is fairly close to the nominal level among all of the situations considered, the estimates ranging between .044 and .057.

WILCOX & CLARK

Negative implications about power seem evident when using TS. As a brief illustration, suppose that data are generated from the model $V = .25X + \varepsilon$, where X and ε are independent and both have a standard normal distribution. Let $Y = 2V$, rounded to the nearest integer. With $n = 60$, power based on TS was estimated to be .073. Using instead HD, power was estimated to be .40.

Table 1. Estimated Type I error probabilities, $\alpha = .05$

Distribution	n	TS	HD
B(3,3,10)	20	0.019	0.044
B(3,3,10)	60	0.002	0.047
B(1,9,10)	20	0.000	0.045
B(1,9,10)	60	0.000	0.045
SN	20	0.011	0.044
SN	60	0.001	0.050
CN	20	0.012	0.057
CN	60	0.004	0.048

Table 2. Estimated Efficiency, $n = 20$

Distribution	TS	TD
SN	0.809	1.090
B(3,3,10)	0.733	0.997
B(1,9,10)	0.689	2.610
CN	2.423	2.487

Of course, when Y has a discrete the least squares estimator could be used. To gain some insight into the relative merits of the HD estimator, its small-sample efficiency was compared to the least squares estimator and the TS estimator for the same situations in Table 1. Let V_0^2 be the estimated squared standard error of least squares estimate of the slope based on 4000 replications. Let V_1^2 and V_2^2 be the estimated squared standard errors for TS and HD, respectively. Then the efficiency associated with TS and HD was estimated with V_0/V_1 and V_0/V_2 , respectively, the ratio of the estimated standard errors. Table 2 summarizes the results. As can be seen, the HD estimator competes very well with the least squares estimator. Moreover, there is no indication that TS ever

ROBUST REGRESSION ESTIMATORS WHEN THERE ARE TIED VALUES

offers much of an advantage over HD, but HD does offer a distinct advantage over TS in some situations.

A related issue is the efficiency of the HD estimator when dealing with a continuous error term, including situations where there is heteroscedasticity. To address this issue, additional simulations were run by generating data from the model $Y = \lambda(X)\varepsilon$ where ε is some random variable having median zero and the function $\lambda(X)$ is used to model heteroscedasticity. The error term was taken to have one of four distributions: normal, symmetric with heavy tails, asymmetric with light tails and asymmetric with heavy tails. More precisely, the error term was taken to have a g-and-h distribution (Hoaglin, 1985) that contains the standard normal distribution as a special case. If Z has a standard normal distribution, then

$$W = \frac{\exp(gZ)-1}{g} \exp(hZ^2), \text{if } g > 0$$

and

$$W = Z \exp\left(h \frac{Z^2}{2}\right), \text{if } g = 0$$

has a g-and-h distribution where g and h are parameters that determine the first four moments. As is evident, $g = h = 0$ corresponds to a standard normal distribution. Table 3 indicates the skewness (κ_1) and kurtosis (κ_2) of the four distributions that were used.

Table 3. Some properties of the g-and-h distribution

g	h	κ_1	κ_2
0.00	0.00	0.00	3.00
0.00	0.20	0.00	21.46
0.20	0.00	0.61	3.68
0.20	0.20	2.81	155.98

WILCOX & CLARK

Three choices for λ were used: $\lambda(X)=1$ (homoscedasticity), $\lambda(X)=|X|+1$ and $\lambda(X)=1/(|X|+1)$. For convenience, these three choices are denoted by variance patterns (VP) 1, 2, and 3.

Table 4 reports the estimated efficiency of TS and HD when X has a normal distribution. To provide a broader perspective, included are the estimated efficiencies of Yohai's (1987) MM-estimator and the least trimmed squares (LTS) estimator. Yohai's estimator was chosen because it has excellent theoretical properties. It has the highest possible breakdown point, .5, and it plays a central role in the robust methods discussed by Heritier et al. (2009). Both the MM-estimator and the LTS estimator were applied via the R package robustbase. As can be seen, for the continuous case, there is little separating the TS, HD and MM estimators with TS and MM providing a slight advantage over HD.

Table 4. Estimated efficiencies, the continuous case, X normal

g	h	VP	TS	HD	MM	LTS
0.000	0.000	1.000	0.861	0.930	0.967	0.708
		2.000	0.994	0.991	1.019	0.769
		3.000	0.997	0.966	0.999	0.776
0.000	0.200	1.000	1.234	1.157	1.199	0.971
		2.000	1.405	1.230	1.267	1.070
		3.000	1.389	1.216	1.276	1.041
0.200	0.000	1.000	0.897	1.146	0.960	0.989
		2.000	1.019	1.009	1.051	0.815
		3.000	0.978	0.999	1.026	0.793
0.200	0.200	1.000	1.314	1.200	1.259	1.022
		2.000	1.615	1.440	1.475	1.197
		3.000	1.443	1.271	1.337	1.160

There are situations where the differences in efficiency are more striking than those reported in Table 4. Also, no single estimator dominates in terms of efficiency: situations can be constructed where each estimator performs better than the others considered here. Suppose, for example, that X has a contaminated normal distribution and Y has a normal distribution. From basic principles, this situation favors OLS because as the distribution of X moves toward a heavy-tailed distribution, the standard error of the OLS estimator decreases. The resulting

ROBUST REGRESSION ESTIMATORS WHEN THERE ARE TIED VALUES

efficiencies were estimated to be 0.514, 0.798, 0.844 and 0.533 for TS, HD, MM and LTS, respectively, with TS and LTS being the least satisfactory. Removing leverage points (outliers among the independent variable) using the MAD-median rule (e.g., Wilcox, 2012a, section 3.13.4), the estimates are 1.336, 1.727, 1.613 and 2.1213. So now LTS performs best in contrast to all of the other situations previously reported.

There is the issue of whether the MM-estimator has good efficiency for the discrete case. For the beta-binomial distribution with $r = s = 3$, the efficiency of the HD estimator is a bit better, but for the other discrete distributions considered here, the efficiency of the MM-estimator could not be estimated because the R function used to compute the MM-estimator routinely terminated with an error. For the same reason, the Type I error probability based on the hypothesis testing method used by the R package `robustbase` could not be studied. Switching to the bootstrap method used here only makes matters worse: bootstrap samples result in situations where the MM-estimator cannot be computed.

An Illustration

Using data from the Well Elderly II study (Jackson et al., 2009), it is illustrated that the choice between the TS and HD estimators can make a practical difference. A general goal in the Well Elderly II study was to assess the efficacy of an intervention strategy aimed at improving the physical and emotional health of older adults. A portion of the study was aimed at understanding the association between the cortisol awakening response (CAR), which is defined as the change in cortisol concentration that occurs during the first hour after waking from sleep, and a measure of Perceived Control (PC), which is the sum of 8 four-point Likert scales. So the possible PC scores range between 8 and 32. Higher PC scores reflect greater perceived control. (For a detailed study of this measure of perceived control, see Eizenman et al., 1997.) CAR is taken to be the cortisol level upon awakening minus the level of cortisol 30-60 minutes after awakening.) Approximately 8% of the PC scores are flagged as outliers using the MAD-median rule. Extant studies (e.g., Clow et al., 2004; Chida & Steptoe, 2009) indicate that various forms of stress are associated with the CAR. After intervention, the TS estimate of the slope is -0.72 with a p-value of .34. Using instead HD, the estimate of the slope is -0.73 with a p-value less than .001.

Concluding Remarks

In summary, when dealing with tied values among the dependent variable, several robust estimators can result in poor control over the Type I error probability and relatively low power, so they should be used with caution. Moreover, the performance of the Theil–Sen estimator can actually deteriorate as the sample size increases. One way of dealing with this problem is to use the HD estimator, which is simple modification of the Theil–Sen estimator. In some situations the HD estimator has better efficiency than other robust estimators, but situations are encountered where the reverse is true. The very presence of tied values does not necessarily mean that robust estimators other than HD will perform poorly. The only point is that when dealing with tied values, the HD estimator can be computed in situations where other robust estimators cannot and it can provide a practical advantage in terms of both Type I error probabilities and power.

Various suggestions have been made about how to extend the Theil–Sen estimator to more than one independent variable (Wilcox, 2012b). One approach is the back-fitting algorithm, which is readily used in conjunction with the HD estimator. Here, the details are not of direct relevance so for brevity they are not provided. An R function, `tshdreg`, has been added to the R package WRS that performs the calculations.

References

Akritas, M. G., Murphy, S. A. & LaValley, M. P. (1995). The Theil–Sen estimator with doubly censored data and applications to astronomy. *Journal of the American Statistical Association* 90, 170-177.

Agostinelli, C. & Markatou, M. (1998) A one-step robust estimator for regression based on the weighted likelihood reweighting scheme. *Statistics Probability Letters*, 37, 341-350.

Chida, Y. & Steptoe, A. (2009). Cortisol awakening response and psychosocial factors: A systematic review and meta-analysis. *Biological Psychology*, 80, 265-278 .

Clow, A., Thorn, L., Evans, P. & Hucklebridge, F. (2004). The awakening cortisol response: Methodological issues and significance. *Stress*, 7, 29-37.

Coakley, C. W. & Hettmansperger, T. P. (1993). A bounded influence, high breakdown, efficient regression estimator. *Journal of the American Statistical Association*, 88, 872-880.

ROBUST REGRESSION ESTIMATORS WHEN THERE ARE TIED VALUES

Dielman, T., Lowry, C. & Pfaffenberger, R. (1994). A comparison of quantile estimators. *Communications in Statistics–Simulation and Computation*, 23, 355-371.

Dietz, E. J. (1987). A comparison of robust estimators in simple linear regression. *Communications in Statistics–Simulation and Computation*, 16, 1209-1227.

Dietz, E. J. (1989). Teaching regression in a nonparametric statistics course. *American Statistician*, 43, 35-40.

Eizenman, D. R., Nesselroade, J. R., Featherman, D. L. & Rowe, J. W. (1997). Intraindividual variability in perceived control in an older sample: The MacArthur successful aging studies. *Psychology and Aging*, 12, 489–502.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. & Stahel, W. A. (1986). *Robust Statistics*. New York: Wiley.

Fernandes, R. & Leblanc, S. G. (2005). Parametric (modified least squares) and non-parametric (Theil-Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment*, 95, 303-316.

Harrell, F. E. & Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika*, 69, 635-640.

Heritier, S., Cantoni, E., Copt, S. & Victoria-Feser, M.-P. (2009). *Robust Methods in Biostatistics*. New York: Wiley.

Hoaglin, D. C. (1985). Summarizing shape numerically: The g-and-h distribution. In D. Hoaglin, F. Mosteller & J. Tukey (Eds.) *Exploring Data Tables Trends and Shapes*. New York: Wiley, pp. 461-515.

Huber, P. J. & Ronchetti, E. (2009). *Robust Statistics*, 2nd Ed. New York: Wiley.

Jackson, J., Mandel, D., Blanchard, J., Carlson, M., Cherry, B., Azen, S., Chou, C.-P., Jordan-Marsh, M., Forman, T., White, B., Granger, D., Knight, B., & Clark, F. (2009). Confronting challenges in intervention research with ethnically diverse older adults: the USC Well Elderly II trial. *Clinical Trials*, 6, 90-101.

Jaekel, L. A. (1972). Estimating regression coefficients by minimizing the dispersion of residuals. *Annals of Mathematical Statistics*, 43, 1449-1458.

Koenker, R. & Bassett, G. (1978). Regression quantiles. *Econometrika*, 46, 33-50.

Liu, R. G. & Singh, K. (1997). Notions of limiting P values based on data depth and bootstrap. *Journal of the American Statistical Association*, 92, 266-277.

Maronna, R. A., Martin, D. R. & Yohai, V. J. (2006). *Robust Statistics: Theory and Methods*. New York: Wiley.

Parrish, R. S. (1990). Comparison of quantile estimators in normal sampling. *Biometrics*, 46, 247-257.

Peng, H., Wang, S. & Wang, X. (2008). Consistency and asymptotic distribution of the Theil-Sen estimator. *Journal of Statistical Planning and Inference*, 138, 1836-1850.

Racine, J. & MacKinnon, J. G. (2007). Simulation-based tests than can use any number of simulations. *Communications in Statistics-Simulation and Computation*, 36, 357-365.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, 79, 871-880.

Sen, P. K. (1968). Estimate of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63, 1379-1389.

Sfakianakis, M. E. & Verginis, D. G. (2006). A new family of nonparametric quantile estimators. *Communications in Statistics-Simulation and Computation*, 37, 337-345.

Sheather, S. J. & Marron, J. S. (1990). Kernel quantile estimators. *Journal of the American Statistical Association*, 85, 410-416.

Staudte, R. G. & Sheather, S. J. (1990). *Robust Estimation and Testing*. New York: Wiley.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Indagationes Mathematicae*, 12, 85-91.

Wang, X. Q. (2005). Asymptotics of the Theil-Sen estimator in simple linear regression models with a random covariate. *Nonparametric Statistics* 17, 107-120.

Wilcox, R. R. (1998). Simulation results on extensions of the Theil-Sen regression estimator. *Communications in Statistics-Simulation and Computation*, 27, 1117-1126.

Wilcox, R. R. (2012a). *Modern Statistics for the Social and Behavioral Sciences: A Practical Introduction*. New York: Chapman Hall/CRC press.

Wilcox, R. R. (2012b). *Introduction to Robust Estimation and Hypothesis Testing*, 3rd Edition. San Diego, CA: Academic Press.

Wilcox, R. R., Erceg-Hurn, D., Clark, F. Carlson, M. (2013). Comparing two independent groups via the lower and upper quantiles. *Journal of Statistical Computation and Simulation*. DOI: 10.1080/00949655.2012.754026

ROBUST REGRESSION ESTIMATORS WHEN THERE ARE TIED VALUES

Yohai, V. J. (1987). High breakdown point and high efficiency robust estimates for regression. *Annals of Statistics*, 15, 642-656.