Scientific
Research
Publishing

A Hausman Type Test for Differences between Least Squares and Robust Time Series Factor Model Betas

Tatiana A. Maravina¹, R. Douglas Martin²

¹University of Washington, Seattle, WA, USA
²Departments of Applied Mathematics and Statistics, University of Washington, Seattle, WA, USA
Email: doug@amath.washington.edu

Abstract

Robust regression is playing an increasingly important role in fitting time series and cross-section factor models for stock returns. We introduce and study the properties of a Hausman type test for comparing factor model regression coefficients computed with LS, which is fully efficient under idealized normal data distributions, and a Robust MM-estimate, which is highly efficient for normally distributed data but also controls variance inflation and bias for outlier generating non-normal data distributions. The test is based on the asymptotic distribution of the difference between the two estimators, one of which is fully efficient. The test can detect a significant difference between the LS and Robust estimate due to the inefficiency of the LS estimator under outlier-generating non-normal error distributions, and due to bias of the LS estimator relative to the Robust estimator caused by bias inducing distributions. The applications efficacy of the new test is demonstrated for comparison of LS and Robust estimates of both CAPM betas and Fama-French three-factor model betas. Monte Carlo studies of the finite sample level and power of the test reveal good performance for sample sizes of at least 100 to 200, which are typical for weekly and daily returns for such models.

Keywords

MM-Estimator, Estimator Efficiency, Estimator Bias, Test For Bias, Hausman Test

1. Introduction

Factor models have an important role in empirical asset pricing and quantitative

portfolio management research, for which a very large literature exists. Examples in empirical asset pricing include papers by Fama and French [1] [2] [3], Hou *et al.* [4] [5], Feng *et al.* [6], and the overview book by Bali *et al.* [7]. Examples in quantitative portfolio management include significant coverage in books such as Grinold *et al.* [8] and Qian *et al.* [9], and papers such as Menchero and Mitra [10], Menchero and Davis [11], Ding and Martin [12], and Ding *et al.* [13]. The main types of factor models appearing in the literature are cross-section factor models and time series factor models, both of which are specific forms of linear regression models.

Linear regression models in quantitative finance are universally fit using ordinary least squares (LS) estimates of the coefficients or weighted least squares (WLS) estimates. Both LS and WLS estimates are relatively simple, widely available in software packages, and blessed by being the best linear unbiased estimates (BLUE) under standard assumptions. In addition, LS estimates are the best among both linear and nonlinear estimates when the errors are normally distributed. However, asset returns and factors often have quite non-normal distributions, and LS coefficient estimates are quite non-robust toward outliers in that they can be very adversely distorted by even one or a few outliers. In statistical terms, LS estimates can suffer from a substantial loss of efficiency when the errors have a fat-tailed non-normal distribution, in that they can have much larger variances than maximum-likelihood estimates (MLEs) for such non-normal distributions. Furthermore, under some types of deviations from normality LS estimates will be biased, even asymptotically as the sample size goes to infinity.

Fortunately, several robust factor model fitting alternatives to LS estimates exist that suffer relatively little from severe inefficiency and bias. See for example the books by Huber [14], Huber and Ronchetti [15], Hampel *et al.* [16], Rousseeuw *et al.* [17], and Maronna *et al.* [18], and the references therein. See also the papers on robust time-series estimation of CAPM betas by Martin and Simin [19], Bailer *et al.* [20], and the paper on robust cross-section factor models by Martin and Xia [21]. Various types of outlier-robust regression methods are implemented in commercial statistical software programs such as SAS and STATA, and in the open-source R packages robust, robustbase, and RobStatTM that are available on CRAN (https://cran.r-project.org/). Regression M-estimates of one form or another are the most widely used robust regression methods.

Statistical inference methods for robust regression coefficients such as robust t-tests, F-tests, robust R-squared, and robust model selection criteria have been available in the literature for many years, and these are described in Maronna *et al.* [18] and are available in the companion R package RobStatTM. On the other hand, the literature on statistical tests for evaluating the difference between LS and robust regressions fits is minimal. In this regard, we recall Tukey [22] who stated "It is perfectly proper to use both classical and robust/resistant methods routinely, and only worry when they differ enough to matter. But when they differ, you should think hard." This is good advice that leaves open the question of

how much is the "enough" in "when they differ enough", and it is highly desirable to have a reliable test statistic whose rejection region defines "enough".

If such a test statistic has a reliable level and adequate power, then acceptance of an appropriately defined null hypothesis would lead a user who routinely computes both LS and robust regressions to be confident in the LS results. On the other hand, rejection of the null hypothesis would support reliance on the robust regression estimate and associated robust inferences. Unfortunately, there does not at present exist a well-accepted statistical test for determining whether LS and robust regression estimates differ significantly from one another. We propose and study the properties of a viable test that uses the robust regression an MM-estimator that is well known in the robust statistics literature.

Our test statistic is focused on differences between LS and Robust MM-estimator factor model slope coefficients, based on a key idea in the specification tests paper by Hausman [23]. We consider composite null and alternative hypotheses where the null hypothesis is that of a linear regression factor model with errors that are normally distributed. The alternative hypothesis consists of outlier generating non-normal error distributions as well as more general types of bias-inducing joint distributions for the returns and factor variables. Rejection of the null hypothesis can occur due to any of the following LS estimator behaviors: inefficiency only, bias only, or both inefficiency and bias.

The novelty of our results is that for the first time there is a reliable significance test for differences between LS and Robust estimates of time series and cross-section factor model coefficients, for selected subsets of coefficients as well as the set of all coefficients. In particular, rejection of the null hypothesis that the data is normally distributed will lead the analyst or risk manager to favor the use of the Robust estimator model fit for risk and performance analysis, and to carry out further analysis to determine the extent and type of non-normality that gives rise to the rejection of the null hypothesis.

2. Robust Regression MM-Estimates

We consider estimation in a linear regression time-series factor model of the form

$$y_t = \boldsymbol{x}^{*\prime}\boldsymbol{\theta} + \epsilon_t = \left(\boldsymbol{1}, \boldsymbol{x}_t'\right)\begin{pmatrix}\alpha\\\boldsymbol{\beta}\end{pmatrix} + \epsilon_t, \quad t = 1, \cdots, N \tag{1}$$

with the assumption that the observed data $z_t = \left(y_t, \boldsymbol{x}_t'\right), t = 1, \cdots, N$, consists of independent and identically distributed random variables. Here, y_t is a return of a specific asset at time t, typically in excess of a risk-free rate, $\boldsymbol{x}_t' = \left(x_{1,t}, x_{2,t}, \cdots, x_{K,t}\right)$ is a vector of K factor returns at time t, α is an unknown intercept, $\boldsymbol{\beta}$ is a K-dimensional vector of unknown regression slope coefficients, and the ϵ_t are the regression errors.

Major applications of such time series factor models in finance include:

- The CAPM model with $K = 1$, where $y_t = r_t$ is an asset return in excess of a risk-free rate at time t, and $x_{1,t}$ is a market return in excess of a risk-free rate at time t,

- The Fama and French [2] 3-factor model (FF3) with $K = 3$, which in addition to the CAPM term, has two more terms: $x_{2,t}$ is a small-minus-big (SMB) factor return, and $x_{3,t}$ is a high-minus-low (HML) factor return,
- The Fama-French-Carhart [24] 4-factor model (FFC4) adds the momentum factor to the FF3 model.

We focus on the important class of robust regression MM-estimators introduced and studied by Yohai [25], which have both the highest possible breakdown point (BP) of 0.5 and high *efficiency* at normal distributions. Efficiency here is defined as the ratio of the variance of the LS estimator to the robust estimator when errors are normally distributed. Since LS has the minimum possible variance at a normal distribution, the efficiency of an MM-estimator, expressed as a percent, is less than 100%. Typically, an efficiency of 85% to 95% is considered high. A regression MM-estimate of $\boldsymbol{\theta}$ is obtained by first computing a high-breakdown point but relatively inefficient initial estimate $\hat{\boldsymbol{\theta}}_0$ and then computing a final estimate $\hat{\boldsymbol{\theta}}$ as the nearest local minimum of

$$\sum_{t=1}^{N} \rho_c \left(\frac{y_t - \boldsymbol{x}_t^{*\prime} \boldsymbol{\theta}}{\hat{\sigma}} \right) \tag{2}$$

With respect to $\boldsymbol{\theta}$, where $\hat{\sigma}$ is a highly robust scale estimate of the residuals. The parameter c is a tuning parameter used to control the trade-off between a high normal distribution efficiency of the estimate and robustness toward outliers, which we discuss subsequently. With $\psi_c = \rho_c'$ the resulting $\hat{\boldsymbol{\theta}}$ satisfies the stationary local minimum condition[1].

$$\sum_{t=1}^{N} \boldsymbol{x}_t^{*} \psi_c \left(\frac{y_t - \boldsymbol{x}_t^{*\prime} \hat{\boldsymbol{\theta}}}{\hat{\sigma}} \right) = \boldsymbol{0} \tag{3}$$

A well-established method of computing $\hat{\sigma}$ and solving the minimization problem (2) was developed by Yohai *et al.* [26], and is briefly described in Appendix B for the interested reader. See also Section 5.5. in Maronna *et al.* [18].

Martin *et al.* [27] demonstrated that to obtain bias-robustness toward outliers, one needs to use a bounded loss function ρ_c. The most popular choice of a bounded loss function is the well-known Tukey bisquare function, and the analytic expressions for the bisquare ρ- and ψ- functions are given in Appendix A. Versions of the bisquare loss functions for normal distribution efficiencies of 85%, 90%, 95%, and 99% are shown on the left plot in Figure 1, and the corresponding psi-functions are on the right. These are all of the rejection type, *i.e.*, they have values of zero outside central regions $(-c, c)$, where the choice of rejection points $\pm c$ determines the normal distribution efficiency.

For reader convenience, the values of constant c and corresponding fractions of data rejected under normality by bisquare psi-function are listed in Table 1 for the four efficiencies of 85%, 90%, 95%, and 99%.

---

[1]Throughout this paper, a prime on a scalar-valued function, e.g. ρ_c', denotes its derivative, otherwise a prime denotes the transpose of a vector or matrix.
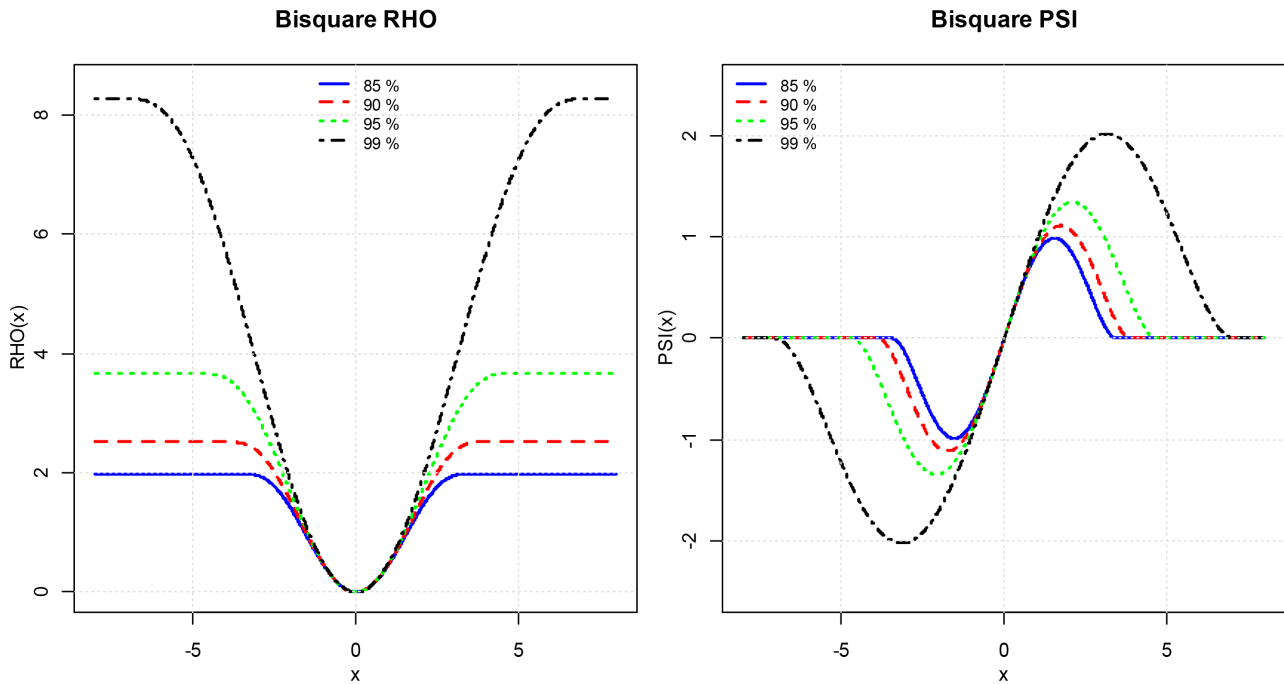
Figure 1. Bisquare rho (left) and psi (right) functions for four normal distribution efficiencies.

Table 1. c values and fractions of data rejected under normality for the bisquare loss functions for four normal distribution efficiencies. The c values are obtained via lmRob.effvy function from R robust package. An observation is said to be rejected when the corresponding psi function is equal to zero. Hence, the fractions of data rejected is $P\left(\left|N\left(0,1\right)\right| > c\right)$.

	Efficiency	85%	90%	95%	99%
Bisquare	c	3.44	3.88	4.68	7.04
	outliers fraction (%)	0.058%	0.010%	0.0003%	$2 \cdot 10^{-10}$%

Consistency and asymptotic normality of MM-estimator was established by [25] under the assumption of independence of the ϵ_t and x_t, where the data $z_t = \left(y_t, x_t'\right), t = 1, \cdots, N$, have a joint distribution

$$F_0\left(x, y\right) = G\left(x\right) F_\epsilon\left(y - x'\beta\right) \tag{4}$$

where x has a finite positive definite covariance matrix C_x. See also Chapter 10 in [18]. In what follows we focus on the LS and MM-estimators of the slopes vector β in (1).

Under model (4) the asymptotic covariance matrix of the LS estimator is

$$V_{LS} = var\left(\epsilon\right) C_x^{-1} \tag{5}$$

and the asymptotic covariance matrix of the MM-estimator $\hat{\beta}_{MM}$ is

$$V_{MM} = \sigma^2 \tau C_x^{-1} \tag{6}$$

$$\tau = \frac{E_{F_\epsilon} \psi_c^2\left(\epsilon/\sigma\right)}{\left(E_{F_\epsilon} \psi_c'\left(\epsilon/\sigma\right)\right)^2} \tag{7}$$

where σ is the asymptotic value of the robust scale $\hat{\sigma}$ estimator.

Under normality the robust scale estimator $\hat{\sigma}$ converges to the standard deviation of the error term, *i.e.* $\sigma^2 = var(\epsilon)$, so that the asymptotic covariance matrices of the MM and LS estimators differ only by the scalar factor τ. This leads to the following convenient relationship under normality.

$$\boldsymbol{V}_{LS} = EFF \cdot \boldsymbol{V}_{MM}. \tag{8}$$

where *EFF* is the large sample normal distribution efficiency of the MM-estimator equal to

$$EFF = \tau^{-1} \tag{9}$$

with ϵ normally distributed with mean 0 and standard deviation σ.

A finite-sample approximation to the covariance matrix of $\hat{\boldsymbol{\beta}}_{MM}$ is obtained by computing estimates of τ, σ and $\boldsymbol{C}_x$. We use a method of doing so proposed by Yohai *et al.* [26] that is described in Sections 5.5 and 5.6 of Maronna *et al.* [18], and implemented in the function lmRob in the R robust library. A brief summary of the method is given in Appendix B.

Here we discuss the behavior of the ordinary least squares (LS) and robust MM estimators under several distinct situations with respect to the joint distribution of the data. First of all, when F_ϵ in model (4) is a normal distribution LS is consistent and fully efficient, and MM is consistent with high efficiency that can be set by the user, e.g., use of 90% or 95% normal distribution efficiency is common. Second, when F_ϵ in model (4) is a non-normal distribution with fat tails but finite variance, the LS estimator is consistent but can have an efficiency arbitrarily close to zero, and the MM-estimators are consistent and can have high efficiencies.

A common approach in robustness studies to allow for more general types of $(y_t, \boldsymbol{x}_t)$ outliers, than those generated by model (4) with a fat-tailed error distribution, is to use a broad family of mixture distributions

$$F(\boldsymbol{x}, y) = (1 - \gamma) F_0(\boldsymbol{x}, y) + \gamma H(\boldsymbol{x}, y) \tag{10}$$

where F_0 is given by (4), the mixing parameter γ is positive and often small, e.g., in the range 0.01 to 0.1, and *H* is unrestricted. This family of models is motivated by the empirical evidence that most of the time the data are generated by the nominal distribution F_0 but with small probability γ the data come from another distribution *H* that can generate a wide variety of outlier types. In the context of the distribution model (10), the goal is to obtain good estimates of the parameters $\theta_0' = (\alpha_0, \boldsymbol{\beta}_0')$ of F_0. Unfortunately, the LS estimator of θ can be not only highly inefficient but also highly biased for some outlier generating distributions *H*. Modern robust regression MM-estimators have been designed to minimize the maximum bias due to unrestricted distributions *H* in the data distribution model (10), while also obtaining high efficiency when $\gamma = 0$ in (10).

3. Test Statistic

The test is designed to test a null hypothesis of a regression model (4) with normally distributed errors. The test is expected to reject when the difference be-

tween LS and robust MM coefficient estimates is largely due to the inefficiency of LS under non-normal error distributions, or due to LS having a large bias relative to the small bias of the MM estimator under the bias-generating data distribution model (10).

It is convenient to motivate our proposed test statistic in the context of estimating the slope parameter β in a simple CAPM linear regression model. Let $\hat{\beta}_{LS}$ and $\hat{\beta}_{MM}$ be LS and robust MM-estimates of β in finite sample sizes. It is known that the efficiency of an inefficient estimate is equal to the squared correlation between the inefficient estimate and an efficient estimate (see for example [28], Theorem 4.8). Thus for $\hat{\beta}_{LS}$ and $\hat{\beta}_{MM}$ under the normality assumption we have:

$$\rho_{MM,LS}^2 = EFF = \frac{Var(\hat{\beta}_{LS})}{Var(\hat{\beta}_{MM})} \tag{11}$$

It follows that under normality:

$$
\begin{aligned}
Var(\hat{\beta}_{LS} - \hat{\beta}_{MM}) &= Var(\hat{\beta}_{LS}) - 2cov(\hat{\beta}_{LS}, \hat{\beta}_{MM}) + Var(\hat{\beta}_{MM}) \\
&= Var(\hat{\beta}_{LS}) - 2\rho_{MM,LS}\sqrt{Var(\hat{\beta}_{LS})Var(\hat{\beta}_{MM})} + Var(\hat{\beta}_{MM}) \\
&= \rho_{MM,LS}^2 Var(\hat{\beta}_{MM}) - 2\rho_{MM,LS}\sqrt{\rho_{MM,LS}^2 Var(\hat{\beta}_{MM})Var(\hat{\beta}_{MM})} + Var(\hat{\beta}_{MM}) \\
&= \rho_{MM,LS}^2 Var(\hat{\beta}_{MM}) - 2\rho_{MM,LS}^2 Var(\hat{\beta}_{MM}) + Var(\hat{\beta}_{MM}) \\
&= (1 - \rho_{MM,LS}^2)Var(\hat{\beta}_{MM}) = (1 - EFF)Var(\hat{\beta}_{MM})
\end{aligned}
$$

In view of (11) the above expression may be written in the following alternative form:

$$Var(\hat{\beta}_{LS} - \hat{\beta}_{MM}) = Var(\hat{\beta}_{MM}) - Var(\hat{\beta}_{LS})$$

A multi-parameter large-sample version of the above result was obtained by Hausman [23] in his classic paper on specification tests[2]. Hausman's Corollary 2.6 to Lemma 2.1 states that the asymptotic covariance matrix of the difference between two consistent and asymptotically normal estimators, one of which is asymptotically fully efficient and the other is inefficient, is equal to the covariance matrix of the inefficient estimator minus the covariance matrix of the efficient estimator. Thus in our case under normality, we have the following asymptotic covariance matrices relationship:

$$Var(\hat{\boldsymbol{\beta}}_{MM} - \hat{\boldsymbol{\beta}}_{LS}) = \boldsymbol{V}_{MM} - \boldsymbol{V}_{LS}. \tag{12}$$

In view of the asymptotic result (8) we have

$$\boldsymbol{V}_{diff} \equiv \boldsymbol{V}_{MM} - \boldsymbol{V}_{LS} = (1 - EFF)\boldsymbol{V}_{MM} = (1 - EFF)\sigma^2 \tau \boldsymbol{C}_x^{-1} \tag{13}$$

Note that (13) holds only under normality because in that case LS is fully efficient and the MM-estimate is inefficient. A result analogous to (12), namely $Var(\hat{\boldsymbol{\beta}}_{MM} - \hat{\boldsymbol{\beta}}_{LS}) = \boldsymbol{V}_{LS} - \boldsymbol{V}_{MM}$, will hold when the MM-estimator is a maximum

---

[2]We thank Professor Eric Zivot for pointing out this reference.

likelihood estimate (MLE) for a non-normal errors distribution and is therefore asymptotically efficient, but the LS estimator is inefficient. Since there is seldom an obvious choice of non-normal distribution MLE to use, we do not pursue this possibility.

Hausman [23] showed that asymptotically $V_{MM} - V_{LS}$ is non-negative definite under normality. However, this result does not hold under non-normal errors, and furthermore positive semi-definiteness of the finite sample estimate of the form $\hat{V}_{diff} = \hat{V}_{MM} - \hat{V}_{LS}$ is not guaranteed even under a normal errors distribution. However, since EFF is less than one, the estimate $\hat{V}_{diff} = (1 - EFF)\hat{V}_{MM}$ is positive definite in the usual situation where the estimate $\hat{V}_{MM}$ is positive definite[3].

By combining LS and MM-estimate regression coefficient estimates with a covariance matrix estimate $\hat{V}_{MM}$ and specified normal distribution efficiency *EFF* of the MM-estimator, one can construct two types of test statistics. The results (12) and (13) suggest the following:

1) A joint test statistic for any subset of k coefficient with $2 \le k \le K$ coefficients:

$$
\begin{aligned}
T_k &= \left(\hat{\boldsymbol{\beta}}_{MM}^{(k)} - \hat{\boldsymbol{\beta}}_{LS}^{(k)}\right)' \left(\frac{1 - EFF}{n}\hat{V}_{MM}^{(k)}\right)^{-1} \left(\hat{\boldsymbol{\beta}}_{MM}^{(k)} - \hat{\boldsymbol{\beta}}_{LS}^{(k)}\right) \\
&= \left(\hat{\boldsymbol{\beta}}_{MM}^{(k)} - \hat{\boldsymbol{\beta}}_{LS}^{(k)}\right)' \left(\frac{1 - EFF}{n}\hat{\sigma}^2 \hat{\tau} \hat{C}_{x,(k)}^{-1}\right)^{-1} \left(\hat{\boldsymbol{\beta}}_{MM}^{(k)} - \hat{\boldsymbol{\beta}}_{LS}^{(k)}\right)
\end{aligned}
\tag{14}
$$

where $\hat{\boldsymbol{\beta}}_{MM}^{(k)}$, $\hat{\boldsymbol{\beta}}_{LS}^{(k)}$, $\hat{V}_{MM}^{(k)}$ and $\hat{C}_{x,(k)}^{-1}$ are the corresponding subsets of $\hat{\boldsymbol{\beta}}_{MM}$, $\hat{\boldsymbol{\beta}}_{LS}$, $\hat{V}_{MM}$ and $\hat{C}_x^{-1}$. For $k = K$, the test is a test for the overall model.

2) A test statistic for any individual coefficient:

$$
T_i = \frac{\hat{\beta}_{MM,i} - \hat{\beta}_{LS,i}}{\sqrt{1 - EFF} \cdot se\left(\hat{\beta}_{MM,i}\right)}
\tag{15}
$$

where

$$
se\left(\hat{\beta}_{MM,i}\right) = \sqrt{\frac{1}{n}\hat{\sigma}^2 \hat{\tau} \hat{C}_{x,ii}^{-1}}
\tag{16}
$$

with $\hat{C}_{x,ii}^{-1}$ equal to the i-th diagonal element of $\hat{C}_x^{-1}$.

Under the null hypothesis of normally distributed errors, the statistic T_k will have approximately a chi-squared distribution with k degrees of freedom and the statistic T_i will have approximately a standard normal distribution. The extent to which the use of such an approximation is valid is explored in Section 5.

4. Two Time Series Factor Model Examples

In this section, we present two pairs of empirical examples of using the proposed test statistic T for determining significant differences between classical LS and the robust bisquare MM estimator with 95% normal distribution efficiency,

---

[3]In principle one might also use the estimate $\hat{V}_{diff} = (EFF^{-1} - 1)\hat{V}_{LS}$. While this estimate should result in decent accuracy of level in finite sample sizes, we conjecture that it will result in lower power under non-normal alternatives due to LS estimates having higher variance than MM estimates.

henceforth referred to as the *Robust estimator*. Subsection 4.1 provides the first pair of examples as follows. First, the test T for a difference between LS and Robust CAPM Betas is computed using observed weekly time series data, with the result that the test rejects the null hypothesis of no difference with a p-value that is zero to three digits. Then outliers that were detected by the Robust estimator were removed and the test was repeated, resulting in accepting the null hypothesis with p-value of 0.387. Subsection 4.2 does likewise for the case of fitting a Fama-French 3-factor model (FF3), which first appeared in Fama and French [2], to weekly stock returns. In this case, using the original data, the test T for the overall model again resulted in a p-value that is zero to three digits. However, with outliers detected by the Robust estimator removed, the overall model test T accepts the null hypothesis with a p-value of 0.217. However, for the FF3 model, we also used the test T for each of the three individual factor coefficients, and for the original data, these tests rejected the null hypothesis for the MKT and HML factors, but not the SMB factor, and with the outliers removed these tests accepted the null hypothesis for all three factors.

4.1. Single Factor CAPM Time Series Model

The single factor model Beta of a set of asset returns is the slope coefficient in a regression of the asset returns on market returns, where both returns are in excess of a risk-free rate. Beta plays a central role in the capital asset pricing model (CAPM) [29] and is one of the most widely known and widely used measures of the expected excess return and market risk of an asset. **Figure 2** shows a scatter plot of the Watts Water Technologies Inc. (WTS) stock weekly returns versus the weekly returns of the CRSP (https://www.crsp.org/) value-weighted market index for the 2-year period from January 2007 to December 2008[4]. The red dashed line shows the least-squares fit, and the black solid line shows the Robust estimator fit. The two dotted lines that are parallel to the solid black line define the outliers-rejection region, *i.e.*, all 6 data values plotted as small circles outside that region are declared outliers[5]. Of the six outliers, the one with the most negative influence on the slope of the LS line is the one with a positive WTS return of close to +0.1 and a negative market return of about -0.18. In addition, the cluster of three outliers with slightly positive market returns but negative WTS returns also has a negative influence on the LS slope. Note that in the legend that displays the LS and Robust estimate values 0.93 and 1.53, the numbers in parentheses are classical and robust standard errors (SEs) of the estimators, and the two beta estimates differ by about 6 times the robust standard error value of 0.103.

The standard error (SE) and p-value for the test T for the difference in the two betas are reported in **Table 2**. Recall from Equation (15) that the SE of T is just a

---

[4]The stock returns data used in this paper are from the "Center for Research in Security Prices, LLC".

[5]Outliers here are defined as asset and market return pairs for which the absolute value of the robust bisquare estimator residual exceeds 3 times a robust residual scale estimate.
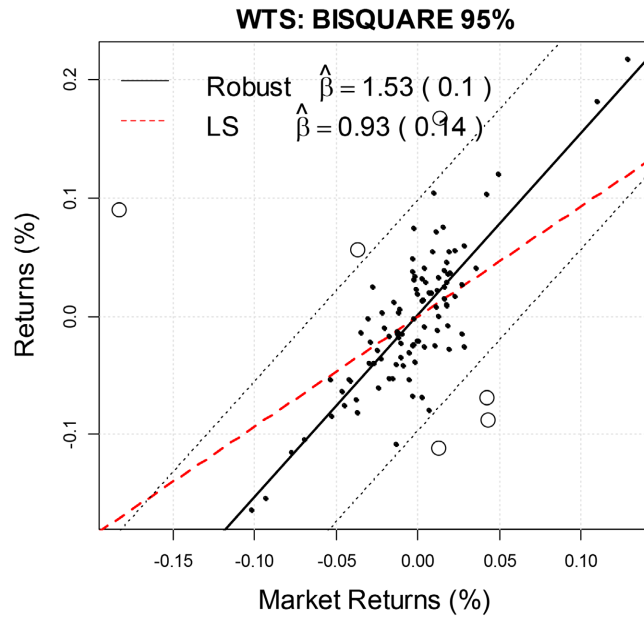
WTS: BISQUARE 95%

Figure 2. Scatter plot of the WTS and market weekly returns in excess of the risk free rate with the fitted LS and Robust lines. The robust residual scale estimate is 0.033.

Table 2. Test statistic for the difference between OLS and robust beta estimates for the WTS example.

$\hat{\beta}_{LS} - \hat{\beta}_{MM}$	SE	p-value
−0.605	0.023	0.000

fraction of the robust beta SE value 0.103, namely $\sqrt{1-0.95} \cdot 0.103 \approx 0.224 \cdot 0.103 = 0.023$, and the corresponding p-value is zero to three digits. The Robust estimator beta value of 1.5 much better describes the stock and market return relationship for the majority of the data than the LS beta value of 0.93. It should be noted that the difference in the two betas of 0.6 would be of considerable financial significance to most investors.

It is interesting to see how the test statistic behaves on a data set that is identical to that of **Figure 2**, except for deleting the 6 outliers of **Figure 2**. The resulting scatter plot shown in **Figure 3** reveals that the LS and Robust estimator coefficients and straight-line fits are now virtually identical. This result illustrates the important characteristic of a good robust fitting method that it gives almost the same results as LS when the data contains no influential outliers, which is also reflected in the high normal distribution efficiency of 95% for the Robust estimator. Not surprisingly then, the **Table 3** p-value of 0.387 indicates no significant difference between the LS and Robust estimates.

Differences between LS and Robust betas are very common, as is revealed in [20]. We highly recommend routine use of robust regression betas along with their standard errors and the test statistic T p-values as a complement to the LS estimates of asset returns provided by many financial data service providers (e.g., ValueLine, Barra, Bloomberg, Capital IQ, Datastream, Ibbotson, Google Finance,
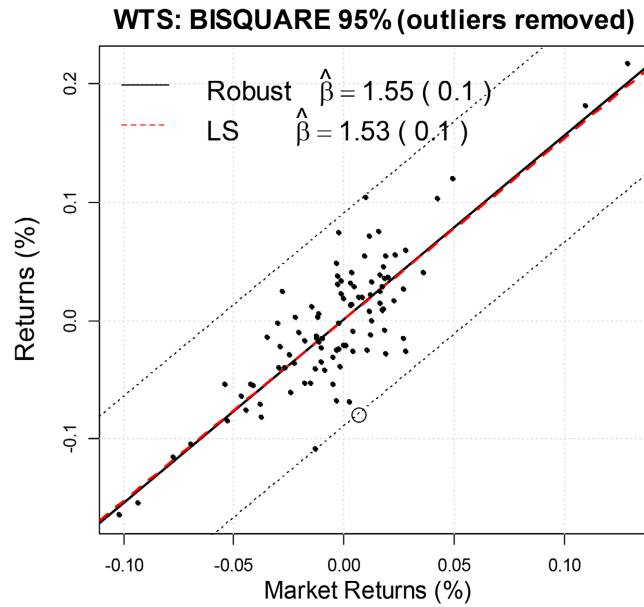
Figure 3. Scatter plot of the WTS and market weekly returns in excess of the risk-free rate with the fitted LS and Robust lines. The robust residual scale estimate is 0.030.

Table 3. Test statistic for the difference between LS and Robust estimates for the WTS example after removing outliers.

$\hat{\beta}_{LS} - \hat{\beta}_{MM}$	SE	p-value
−0.019	0.022	0.387

and others). Acceptance of a null hypotheses of no significant differences between the robust and LS betas would give investors extra comfort in making their decisions based on the classical LS beta estimates. On the other hand, large significant differences should alert analysts to investigate the returns data closely to determine which beta are the most useful guide to investment decisions.

4.2. Multifactor Time Series Model

Here we apply our test T to the LS and Robust fits of the Fama-French 3 factor model (FF3) to the weekly returns of the stock with ticker ADL for the year 2008. The FF3 time series factor model has the form:

$$r_t^e = \alpha + f_{MKT,t}^e \beta_1 + f_{SMB,t} \beta_2 + f_{HML,t} \beta_3 + \epsilon_t \tag{17}$$

where r_t^e is a time series of the asset excess returns relative to a risk-free rate, $f_{MKT,t}^e$ is a time series of market excess returns, $f_{SMB,t}$ are the returns of the Fama-French "small minus big" (SMB) size factor portfolio, $f_{HML,t}$ are the returns on the "high minus low" (HML) value factor portfolio.

The time series of the ADL weekly returns and the FF3 MKT, SIZE, and HML weekly factors for 2008 are shown in **Figure 4**[6]. Recall that by late 2008 the

[6]The stocks and MKT returns are from the "Center for Research in Security Prices, LLC", and the definitions of the SMB and HML factors and their time series of factor returns are available at Professor Ken French's website https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
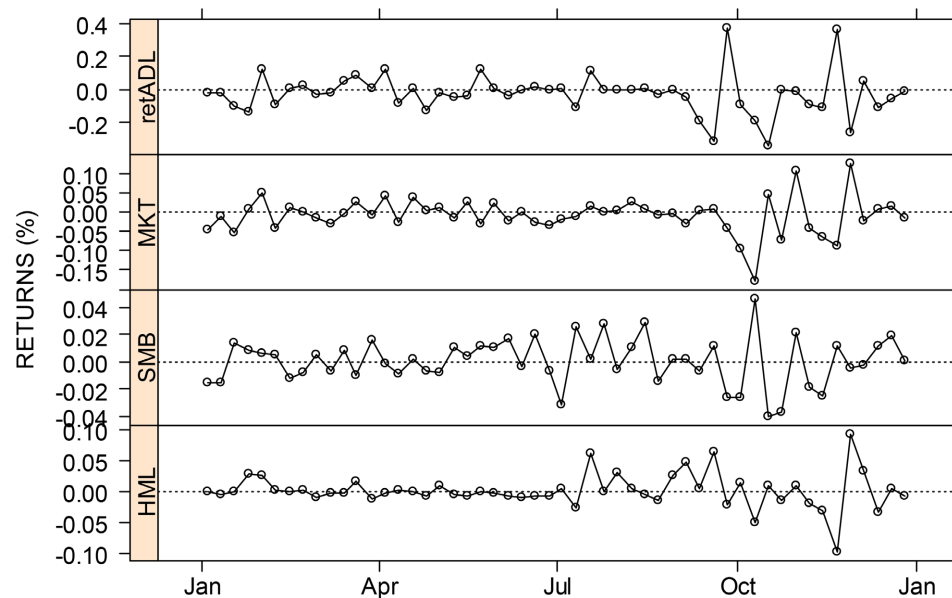
Figure 4. Time series of the ADL 2008 weekly returns and corresponding MKT, SIZE and HML factors.

financial crisis that began in 2007 was in full force, and this is reflected in the increased volatility and outlier values to various degrees and timing across the ADL returns and FF3 factors. Specifically, one sees in **Figure 4** that increased volatility starts for retADL in late September of 2008, while that of MKT starts in early October, and that of the SMB and HML factors starts near the beginning of July 2008. By the end of December 2008, the retADL, MKT, SMB, and HML volatilities have all decreased to pre-crisis levels.

The **Figure 5** display of the pairwise scatter plots of the ADL returns and FF3 factors reveals clear outliers in each panel, and one expects to get a better FF3 model fit to the ADL returns with the Robust estimator than with LS.

Figure 6 contains the time series of appropriately scaled residuals from the LS fit in the left panel, and from the Robust fit in the right panel. For the LS estimator, the residuals are scaled by the errors standard deviation estimate, and for the Robust estimator, the residuals are scaled by a robust scale estimate of the errors. The horizontal dotted lines located at +3 and −3 define a central region outside of which scaled residuals are considered outliers. Clearly, the LS fit gives no warning whatsoever in the time series of standard deviation scaled residuals that the data contains influential outliers, whereas 5 robustly scaled residuals outliers are clearly revealed by the Robust fit.

Figure 7 displays normal QQ-plots of the residuals from the LS and Robust fits. The dashed lines in the plots are the pointwise 95% simulated confidence intervals. Although the LS normal QQ-plot shows some signs of non-normality by virtue of the deviation from the solid straight line in the central region of the plot, no data points fall outside the region defined by the dashed lines, so one sees no evidence of significant deviation from the normality of the residuals from the LS fit. On the other hand, for the Robust fit, the normal QQ-plot not only fits the
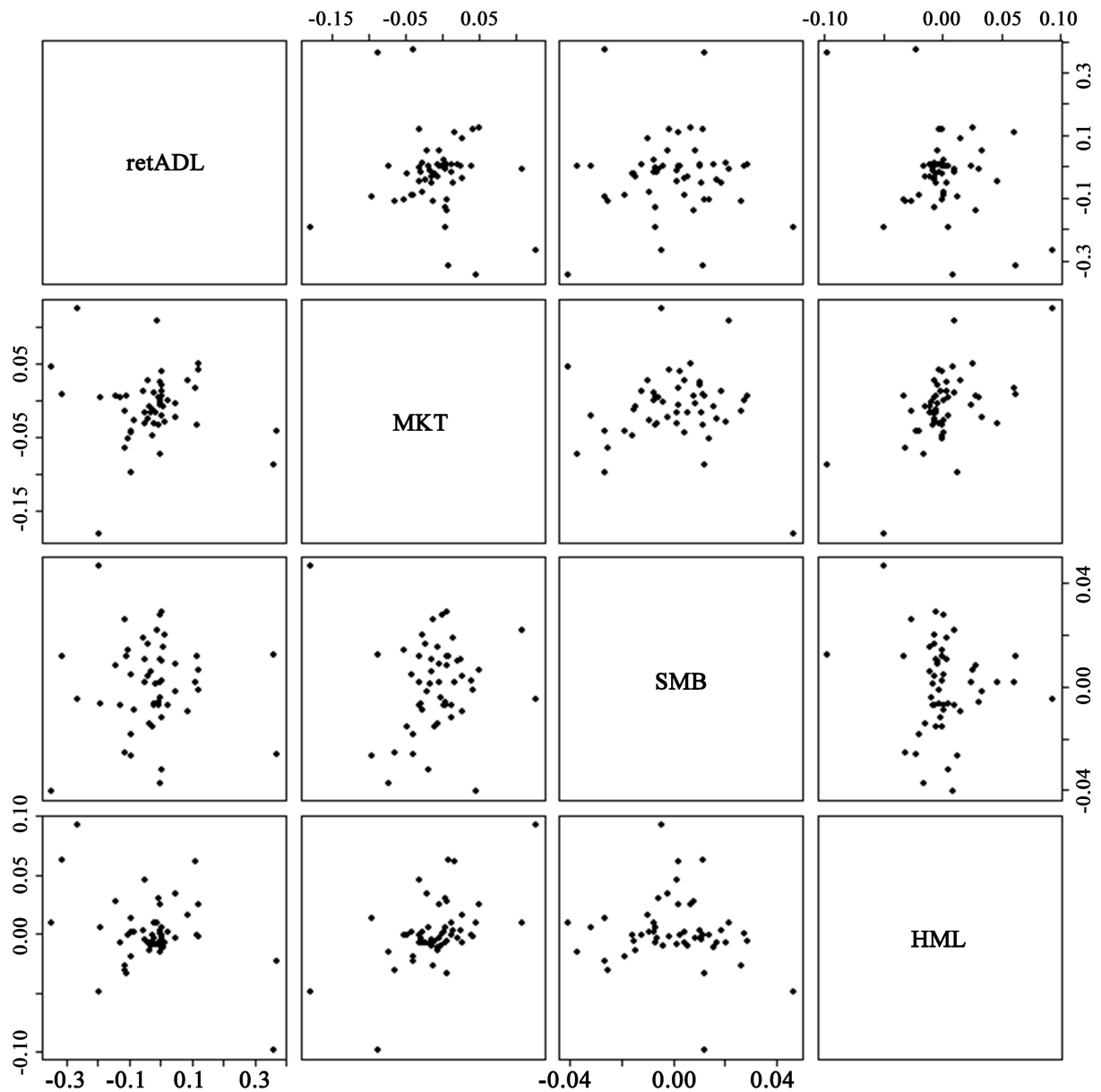
Figure 5. Pairwise scatterplots of ADL 2008 weekly returns and corresponding MKT, SMB and HML factors.

straight line quite well in the central region of the plot, but it also very clearly exposes the 5 outliers that appear in the Robust fit panel of Figure 6, as well as 3 other residuals that fall just outside the region defined by the dashed lines.

The LS and Robust coefficient estimates and their differences based on fitting the FF3 model to the ADL returns are displayed in Table 4, together with their individual standard errors (SE) and p-values, along with the SE's and p-values for the test T of the difference in the two coefficients for each of the three factors. This test rejected the null hypothesis for the MKT and HML factors with a p-value of zero to the three digits. However, the difference between LS and Robust coefficients for the SMB factor is small and is not statistically significant (test T p-value is 0.539). The joint test statistic based on the 3 degrees of freedom

Table 4. Regression results for the ADL multi factor example. The joint test statistic is 1101 on 3 DF, with a p-value = 0.000.

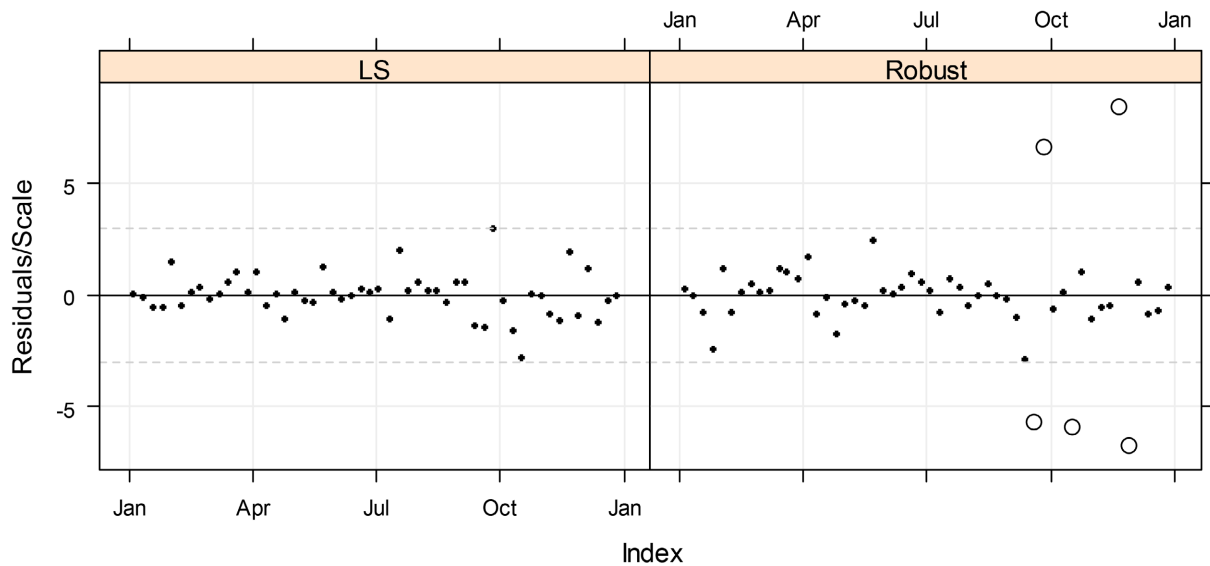|	LS			Robust Estimate			T		
	Coef	SE	P-value	Coef	SE	P-value	$\hat{\beta}_{LS} - \hat{\beta}_{MM}$	SE	P-value
Alpha	−0.021	0.017	0.236	−0.015	0.009	0.117			
MKT	0.341	0.433	0.435	0.652	0.228	0.006	−0.310	0.051	0.000
SMB	−0.272	0.979	0.783	−0.346	0.539	0.524	0.074	0.121	0.539
HML	−1.895	0.727	0.012	1.125	0.503	0.030	−3.020	0.113	0.000

Figure 6. Time series of least squares residuals (LS panel) and robust bisquare residuals (Robust panel).

chi-squared distribution approximation has the value 1101 and a p-value that is zero to 3 digits.

Table 5 reports the same quantities as in Table 4 except that the ADL stock and FF3 factor returns are deleted at the times of the five residuals outliers in the right-hand panel of Figure 6. Not surprisingly, neither the overall test nor the individual coefficients tests reject the null hypothesis of normally distributed linear factor model errors. This is consistent with the high 95% normal distribution efficiency of the Robust estimator, along with the fact that after deleting the 5 most extreme outliers in the right-hand panel of Figure 7, the distribution of the residuals is close to a normal distribution. Note also that the only coefficients that changed substantially after removing outliers were the LS estimates for the MKT and HML factors, the two factors for which original test T in Table 4 indicated a statistically significant difference with the Robust estimates. As expected, the removal of outliers did not affect the Robust estimates much.

5. Monte Carlo Simulations

In order to evaluate the finite sample behavior of the level and power of our test

Table 5. Regression results for the ADL multi factor example after removing outliers. Joint test statistic is 4.45 on 3 DF, with a p-value 0.217.

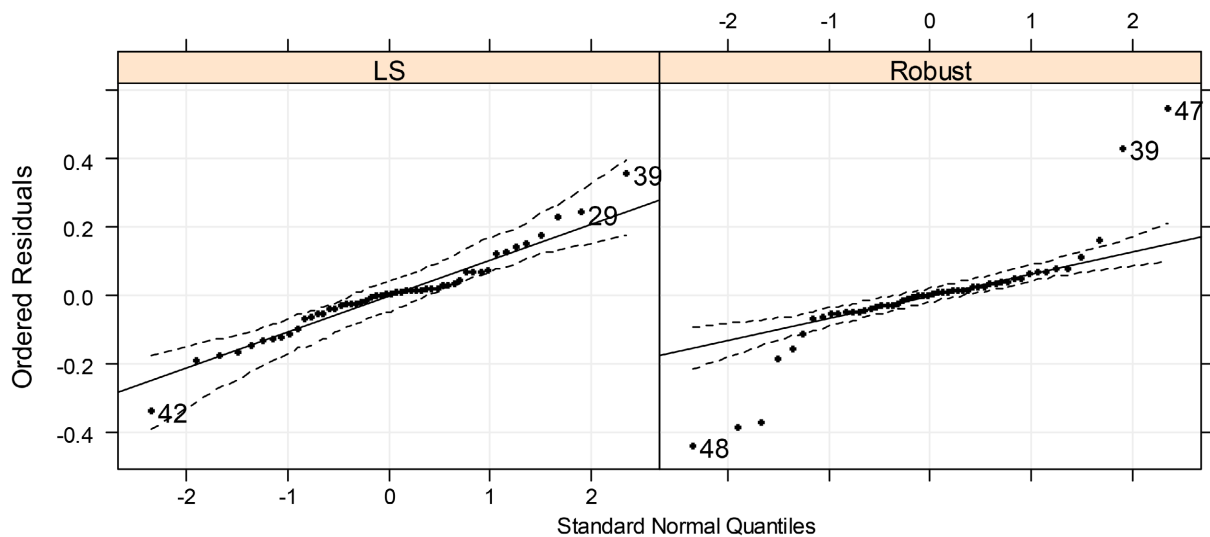|	LS			Robust Estimate			T		
	Coef	SE	P-value	Coef	SE	P-value	$\hat{\beta}_{LS} - \hat{\beta}_{MM}$	SE	P-value
Alpha	−0.017	0.010	0.087	−0.014	0.009	0.138			
MKT	0.625	0.245	0.014	0.659	0.227	0.006	−0.034	0.051	0.503
SMB	−0.270	0.585	0.647	−0.361	0.533	0.502	0.090	0.119	0.449
HML	1.019	0.538	0.065	1.165	0.501	0.025	−0.146	0.112	0.193

Figure 7. Normal QQ-plots of the residuals for the LS fit (left panel) and the robust bisquare fit (right panel).

as a function of the choice of a normal distribution efficiency, we carried out a number of Monte Carlo simulation studies with a large-sample significance level α set at 0.05. As approximations to the finite sample level and power of the tests, we calculate Monte Carlo rejection rates, *i.e.*, the proportion of times out of M replicates that a given hypothesis was rejected. Because of the relative complexity of the analysis, we focus on the slope coefficient β in a single factor model $y_t = \alpha_0 + \beta_0 x_t + \epsilon_t$, $t = 1, \cdots, N$, where under the CAPM the intercept $\alpha_0 = 0$. Simulations were conducted in R (version 2.13.0) using the lmRob function from the R robust library. Note that the test statistic given by (15) is easily computed from the output of lmRob and the standard R least-squares fitting function lm.

5.1. Distribution Models

We assume independent and identically distributed (i.i.d.) random x_t that are independent of i.i.d. errors ϵ_t for the first two models below. We generate samples from the following distributions for the errors ϵ_t:

Model 1: Standard normal, which is included in the null hypothesis

Model 2: Skew-t distribution of Azzalini and Capitanio [30] with skewness

parameter $\lambda = 1$, and 5 degrees of freedom, as implemented in the R package sn, which is included in the composite alternative.

Model 3: Asymmetric two-term conditional joint normal mixture for x_t and ϵ_t that is included in the composite alternative:

$$\begin{pmatrix} x_1 \\ \epsilon_1 \end{pmatrix}, \cdots, \begin{pmatrix} x_N \\ \epsilon_N \end{pmatrix} \text{ are i.i.d. } (1-\gamma) N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) + \gamma N\left(\begin{pmatrix} \mu_x \\ \mu \end{pmatrix}, \begin{bmatrix} 0.25^2 & 0 \\ 0 & 0.25^2 \end{bmatrix}\right) \tag{18}$$

where we condition on the number of "outliers" from the second component to be $\lfloor \gamma N \rfloor$, with γ ranging from 0.01 to 0.1, $\mu_x = 2$, $\mu = 4$ and 7. In this case the mixture model is such that large positive residuals occur for large values of x_t, and result in biased LS and robust MM-estimates of β, with the bias of the latter being much smaller than for LS.

We carry out the conditioning in Model 3 as follows. We first generate $x_1, \cdots, x_N$ as i.i.d. $N(0, 1^2)$ and $\epsilon_1, \cdots, \epsilon_N$ as i.i.d. $N(0, 1^2)$. Then we randomly select $\lfloor \gamma N \rfloor$ observations and replace corresponding ϵ_t with i.i.d. $(\mu \gg 0, 0.25^2)$, and also replace the corresponding x_t with i.i.d. $N(\mu_x, 0.25^2)$. As a result, the reported null hypotheses rejection rates are not confounded by the randomness of the outlier fraction in each sample. The corresponding unconditional rejection rates for these two models can be easily obtained from conditional rejection rates as $\sum_{i=0}^{\infty} RR_i \cdot p_i$, where RR_i is a conditional rejection rate when the number of outliers is equal to i, and p_i is the probability that the number of outliers is equal to i. For example, for $\gamma = 0.02$ about 13.3% ($p_0 = 0.133$) of the samples of size 100 will have no outliers, 27.1% ($p_1 = 0.271$) will have exactly one outlier, 27.3% ($p_2 = 0.273$) will have exactly two outliers, and 32.3% of the samples will have three or more outliers.

For all 3 models we set $\alpha_0 = 0$, $\beta_0 = 1$. For models 1 and 2 we generated 10,000 replicates. Model 3 includes many combinations of the parameters μ and γ, and for each such combination we generated 1000 replicates[7]. We used sample size N ranging from 50 to 500.

5.2. Results

Model 1 (normal distribution errors). **Figure 8** displays the normal distribution Monte Carlo level versus sample size for the Robust estimator, and for the four normal distribution efficiencies (85%, 90%, 95%, and 99%).

The actual level of the test is generally larger than the nominal significance level of 0.05 for all four normal distribution efficiencies, and decreases with increasing sample size, except for the curious constant or increasing actual level

---

[7]Standard errors for the Monte Carlo level and power estimates can be obtained from the estimation theory of a binomial proportion p. In particular, using classical standard errors, namely $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{M}}$, we see that the standard errors are reasonably small even at $M = 1000$ replicates. The standard errors of the Monte Carlo level, i.e. when $\hat{p} = 0.05$, are approx. 0.0069 for $M = 1000$ and 0.0022 for $M = 10000$. The standard error of the Monte Carlo power is the largest when $\hat{p} = 0.5$, and in this case is equal to 0.0158 for $M = 1000$ and equal to 0.005 for $M = 10000$.
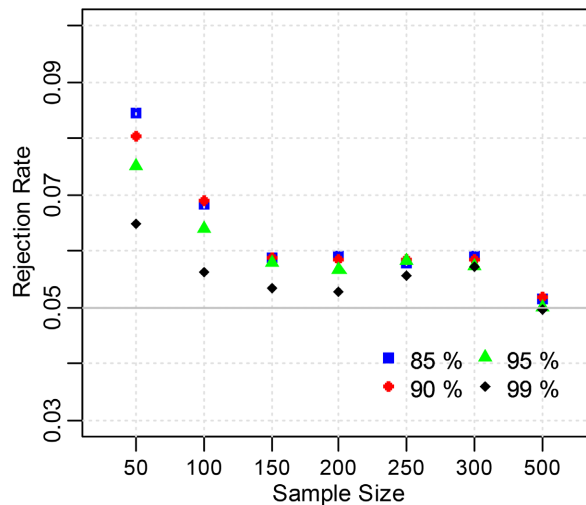
Figure 8. Model 1. Level of the test T for the CAPM β in the single factor model with normal residuals. The grey horizontal line is at a large-sample significance level of 0.05.

for sample sizes 250 and 300, and then being right at 5% at sample size 500. It is striking that the actual levels are uniformly closest to 5% for the Robust estimator with 99% normal distribution efficiency. For estimating a CAPM beta with a sample size of 50 for 4 years and 2 months of monthly returns, the actual error rate for even that best 99% efficiency estimator has an unacceptably high level of about 6.5%. At all larger sample sizes, the level of the 99% efficient Robust estimator levels are all substantially less than 6% and right on the asymptotic level target of 5% at sample size 500. It is notable that for sample sizes 250 and above the Robust estimators at all four normal distribution efficiencies have essentially equivalent test levels. It remains to explore Monte Carlo results with larger numbers of replications to see if the unnatural non-monotonic behavior of the test levels for sample sizes 250 and 300 will disappear.

Model 2 (skew-t distribution errors). Results for a skewed t-distribution with five degrees of freedom are displayed in **Figure 9**.

The skewed t-distribution is in the alternative hypothesis for the test and thus one would hope for high power results. The power indeed increases with increasing sample size and with normal distribution efficiency. It can be shown that the power of the test T for sample size 500 is close to the estimated asymptotic value for each of the four efficiencies. Since both the LS and robust estimates are consistent estimators that converge to beta β_0 at the same rate, the asymptotic power of the test T will be less than one, and it is not surprising that the power of T is less than one for the largest sample sizes in **Figure 9**.

Model 3 (bivariate normal mixture distribution). Under this model, the LS and MM slope estimates converge to different values, and so one anticipates high power of the test for differences between LS and MM estimates, more so the larger the sample size and the larger the μ. The power of the test for $\mu = 4$ and 7 are shown in **Figure 10**, and the reasons for the results are as follows. First, we note that the power is essentially 100% for all sample sizes and all normal distribution
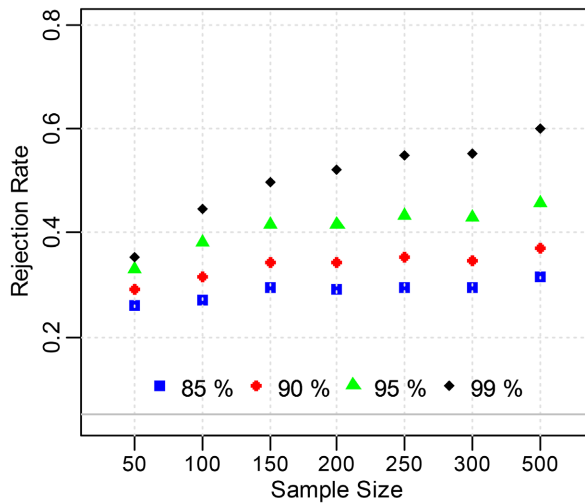
Figure 9. Model 2. Power of the test for the CAPM β in the single factor model for skewed t_5 residuals. The grey horizontal line is at a large-sample significance level of 0.05.
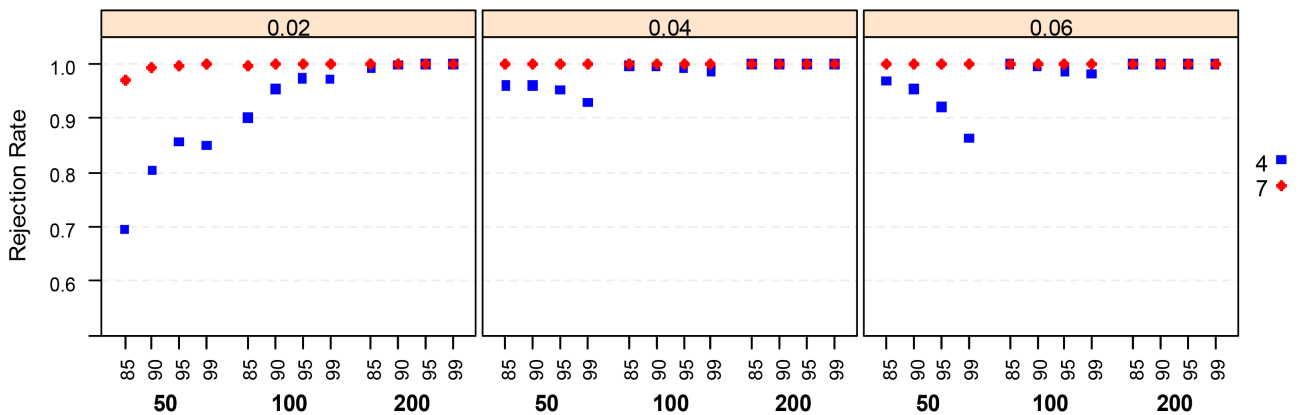

Figure 10. Model 3. Power of the test for the CAPM β in the single factor model under bivariate asymmetric contamination with $\mu_x = 2$. Squares correspond to smallish outliers due to the value $\mu = 4$, and diamonds correspond to large outlies due to $\mu = 7$.

efficiencies for $\mu = 7$. This should not be surprising since in this case the outlier sizes are quite large relative to the central standard normal distribution in (18), and the outliers are rejected by the Robust estimator. The case of $\mu = 4$ is more challenging as the outliers are only smallish and are not rejected and only down-weighted by the Robust estimator (see **Table 1**). Not surprisingly, for each γ and each normal distribution efficiency the power of T increases with increasing sample size. For sample sizes 100 and 200 with $\mu = 4$, the power of T is essentially 100% for $\gamma = 0.04$ and 0.06, and for $\gamma = 0.02$ the power increases to close to 100% as the normal distribution efficiency increases from 85% to 99%. For sample size 50, which is close to the commonly used sample size 60 for estimating a CAPM beta with 5 years of monthly returns, the quite similar decreasing power versus normal distribution efficiency relationship for $\gamma = 0.04$ and 0.06 is distinctly different than the increasing relationship for $\gamma = 0.02$. Thus, there is an interaction between the value of γ and the Robust estimator efficiency, for which we do not yet have a clear explanation.

A surprising result of the Model 1 and Model 2 Monte Carlo study, revealed in **Figure 8** and **Figure 9**, is that more accurate levels and higher power are obtained using the robust MM-estimator with a higher normal distribution efficiency of 99% instead of the more traditional 95%. The reason this is surprising is that using lower normal distribution efficiency of an MM-estimator generally results in lower bias due to bias -inducing outliers. Note however that for sample size 50 in **Figure 10**, higher efficiency yields higher power only for the fraction 0.02 of outliers, and lower efficiency yields higher power for fractions 0.04 and 0.06. This represents a curious interaction between the fraction of outliers and MM-estimator efficiency, and this behavior needs further study.

We remark that the increasing empirical levels of the test T as the sample size decreases in **Figure 4** below sample size 150 is likely due to a small sample bias in the estimate $se\left(\hat{\beta}_{MM,i}\right)$ that appears in the denominator of T in (15). It will be worthwhile to consider possible bias correction methods to improve the small sample size accuracy of the level of the test.

6. Summary and Discussion

This paper uses the important Hausman [23] result to construct a new test statistic T for detecting differences between LS and Robust estimators of the slope coefficients of time series and cross-section factor models. This test is available as the test T1 in the function lsRobTest contained in the robust package on CRAN (https://cran.r-project.org/web/packages/robust/index.html). Rejection of the test supports the use of the Robust estimator instead of, or as a diagnostic complement to, the LS estimator, and investigation of influential returns and factors outliers and their cause. The efficacy of the test statistic T is demonstrated with two different factor model application examples, and by an extensive Monte Carlo study of the level and power of the test. These results support the routine use of the new test statistic T as a mathematically and empirically justified new method of detecting significant differences between LS estimates and Robust MM-estimates of time-series factor models. The test is expected to be equally valuable in the context of cross-section factor models, and indeed for any linear regression model.

A limitation of the proposed test is that rejection of the null hypothesis of no difference between the LS and Robust estimator does not tell us whether rejection occurred due to inefficiency of LS under Model (4) with non-normal errors, or bias of LS being larger than that of the MM estimator under Model (10), or both inefficiency and bias. It is a topic for further research to design a test or tests, that can inform the researcher which of the above deviations from the ideal normal distribution model gives rise to rejection of the null hypothesis.

Finally, last but not least, we have focused on the slopes in the factor Model (1) and have ignored the intercept α, which is often quite important in cross-section and time-series factor models. For example, a test of the null hypothesis that $\alpha = 0$ is important in tests of the validity of the CAPM model. For-

tunately, there is a simple method to take care of this by centering the factor model response and factor exposures with sample medians and using the MM-estimate of regression through the origin for these transformed variables. The robust slope coefficients obtained in this manner can be used to compute regression residuals whose median will be a robust estimate of the intercept. We plan to study the statistical properties of the resulting robust intercept estimate in a separate follow-on study.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Fama, E.F. and French, K.R. (1992) The Cross-Section of Expected Stock Returns. *Journal of Finance*, **47**, 427-465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

[2] Fama, E.F. and French, K.R. (1993) Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, **33**, 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

[3] Fama, E.F. and French, K.R. (2015) A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, **116**, 1-22. https://doi.org/10.1016/j.jfineco.2014.10.010

[4] Hou, K., Xue, C. and Zhang, L. (2020) Replicating Anomalies. *The Review of Financial Studies*, **33**, 2019-2133. https://doi.org/10.1093/rfs/hhy131

[5] Hou, K., Mo, H., Xue, C. and Zhang, L. (2021) An Augmented q-Factor Model with Expected Growth. *Review of Finance*, **25**, 1-41. https://doi.org/10.1093/rof/rfaa004

[6] Feng, G., Giglio, S. and Xiu, D. (2020) Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance*, **75**, 1327-1370. https://doi.org/10.1111/jofi.12883

[7] Bali, T.G., Engle, R.F. and Murray, S. (2016) Empirical Asset Pricing: The Cross Section of Stock Returns, John Wiley & Sons, Hoboken. https://doi.org/10.1002/9781118445112.stat07954

[8] Grinold, R.C. and Kahn, R.N. (2000) Active Portfolio Management: A Quantitative Approach for Providing Superior Returns and Controlling Risk. McGraw-Hill, New York.

[9] Qian, E.E., Sorensen, E.H. and Hua, R.H. (2007) Quantitative Equity Portfolio Management: Modern Techniques and Applications. CRC Press, Boca Raton. https://doi.org/10.1201/9781420010794

[10] Menchero, J. and Mitra, I. (2008) The Structure of Hybrid Factor Models. *Journal of Investment Management*, **6**, 35-47.

[11] Menchero, J. and Davis, B. (2010) The Characteristics of Factor Portfolios. *Journal of Performance Measurement*, **15**, 52-62.

[12] Ding, Z. and Martin, R.D. (2017) The Fundamental Law of Active Management: Redux. *Journal of Empirical Finance*, **43**, 91-114. https://doi.org/10.1016/j.jempfin.2017.05.005

[13] Ding, Z., Martin, R.D. and Yang, C. (2020) Portfolio Turnover When IC Is Time-Varying. *Journal of Asset Management*, **21**, 609-622. https://doi.org/10.1057/s41260-019-00145-1

[14] Huber, P.J. (1981) Robust Statistics. John Wiley & Sons, Inc., New York. https://doi.org/10.1002/0471725250

[15] Huber, P.J. and Ronchetti, E.M. (2009) Robust Statistic. 2nd Edition, John Wiley & Sons, Hoboken. https://doi.org/10.1002/9780470434697

[16] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986) Robust Statistics: The Approach Based on Influence Functions. John Wiley & Sons, Inc., New York.

[17] Rousseeuw, P.J. and Leroy, A.M. (1987) Robust Regression and Outlier Detection. John Wiley & Sons, Inc., New York. https://doi.org/10.1002/0471725382

[18] Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibian-Barerra, M. (2019) Robust Statistics: Theory and Methods (with R). 2nd Edition, Wiley, Hoboken. https://doi.org/10.1002/9781119214656

[19] Martin, R.D. and Simin, T.T. (2003) Outlier-Resistant Estimates of Beta. *Financial Analysts Journal*, **59**, 56-69. https://doi.org/10.2469/faj.v59.n5.2564

[20] Bailer, H.M., Maravina, T.A. and Martin, R.D. (2012) Robust Betas in Asset Management. In: Scherer, B. and Winston, K., Eds., *The Oxford Handbook of Quantitative Asset Management*, Oxford University Press, Oxford, 203-242. https://doi.org/10.1093/oxfordhb/9780199553433.013.0011

[21] Martin, R.D. and Xia, D.Z. (2021) Efficient Bias Robust Cross-Section Factor Models. SSRN, 1-53. https://doi.org/10.2139/ssrn.3921175

[22] Tukey, J.W. (1979) Robust Techniques for the User. In: Launer, R.L. and Wilkinson, G.N., Eds., *Robustness in Statistics*, Academic Press, New York, 103-106. https://doi.org/10.1016/B978-0-12-438150-6.50013-3

[23] Hausman, J.A. (1978) Specification Tests in Econometrics. *Econometrica*, **46**, 1251-1271. https://doi.org/10.2307/1913827

[24] Carhart, M.M. (1997) On Persistence in Mutual Fund Performance. *Journal of Finance*, **52**, 57-82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

[25] Yohai, V.J. (1987) High Breakdown-Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistics*, **15**, 642-656. https://doi.org/10.1214/aos/1176350366

[26] Yohai, V.J., Stahel, W.A. and Zamar, R.H. (1991) A Procedure for Robust Estimation and Inference in Linear Regression. In: Stahel, W. and Weisberg, S., Eds., *Directions in Robust Statistics and Diagnostics* (*Part II*), Springer, New York, 365-374. https://doi.org/10.1007/978-1-4612-4444-8_20

[27] Martin, R.D., Yohai, V.J. and Zamar, R.H. (1989) Min-Max Bias Robust Regression. *The Annals of Statistics*, **17**, 1608-1630. https://doi.org/10.1214/aos/1176347384

[28] Lehmann, E.L. and Casella, G. (1998) Theory of Point Estimation. 2nd Edition, Springer, Berlin.

[29] Sharpe, W.F. (1964) Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, **19**, 425-442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

[30] Azzalini, A. and Capitanio, A. (2003) Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew T Distribution. *Journal of the Royal Statistical Society*: Series B, **65**, 367-389. https://doi.org/10.1111/1467-9868.00391

[31] Rousseeuw, P.J. and Yohai, V.J. (1984) Robust Regression by Means of S-Estimators. In: Härdle, F.W. and Martin, R.D., Eds., *Robust and Nonlinear Time Series*, *Lecture Notes in Statistics* 26, Springer, New York, 256-272. https://doi.org/10.1007/978-1-4615-7821-5_15

[32] Hampel, F.R. (1971) A General Qualitative Definition of Robustness. *The Annals of Mathematical Statistics*, **42**, 1887-1896. https://doi.org/10.1214/aoms/1177693054

Appendix A: Analytical Expressions for Bisquare Functions

The analytic expressions for the bisquare ρ and ψ functions are:

$$\rho_c(r) = \begin{cases} 1 - \left(1 - \left(\dfrac{r}{c}\right)^2\right)^3, & \left|\dfrac{r}{c}\right| \le 1 \\ 1, & \left|\dfrac{r}{c}\right| > 1 \end{cases}$$

and

$$\psi_c(r) = \frac{6}{c^2} r \left(1 - \left(\frac{r}{c}\right)^2\right)^2 I\left[|r| \le c\right]$$

These are the rescaled bisquare ρ and ψ functions that are computed by the functions rho.weight (…, ips = 2) and psi.weight (…, ips = 2) in R robust library. **Figure 1**, however, plots unscaled versions $\dfrac{c^2}{6}\rho_c(r)$ and $\dfrac{c^2}{6}\psi_c(r)$.

Appendix B: Regression MM-Estimates

Consider a regression model (1) and let $\theta' = (\alpha, \beta')$ and $x_t^{*'} = (1, x_t')$ and $p = K + 1$, the length of parameters vector θ. An effective computational procedure for MM-estimates was developed in Yohai *et al*. [26] and consists of the following key steps, where it is assumed that the bisquare function $\rho_c(u)$ is used, and without loss of generality is standardized so that $\max_u \rho_c(u) = 1$:

1) Compute an initial robust S-estimate $\hat{\theta}_1$ with high breakdown point of one-half, but low normal distribution efficiency as follows (see Rousseeuw and Yohai [31]).

With $c_1 = 1.548$ and for any θ let $s_{c_1}(\theta)$ be the solution of

$$\frac{1}{n-p} \sum_{t=1}^{N} \rho_c \left(\frac{y_t - x_t^{*'}\theta}{s_{c_1}(\theta)} \right) = 0.5$$

where the value 0.5 on the right-hand side ensures that the S-estimator $\hat{\theta}_1$ has a BP = 0.5.

The regression S-estimate of θ is a value $\hat{\theta}_1$ that minimizes $s_{c_1}(\theta)$:

$$\hat{\theta}_1 = \arg\min_\theta s_{c_1}(\theta) \tag{A.1}$$

2) Let $\hat{\sigma}_1 = s_{c_1}(\hat{\theta}_1)$ be the robust scale estimate determined by the regression S-estimator $\hat{\theta}_1$. The choice $c_1 = 1.548$ is used so that $\hat{\sigma}_1$ is a consistent estimator of the standard deviation of the ϵ_t when they have a normal distribution.

3) The final estimate $\hat{\theta}_2$ is obtained as a local minimum of (2) with $\hat{\sigma} = \hat{\sigma}_1$, that is nearest to $\hat{\theta}_1$, where the loss function is now ρ_{c_2} with $c_2 > c_1$ chosen to yield a user-specified "high" normal distribution efficiency. Constant c_2 values for the normal efficiencies of 85%, 90%, 95% and 99% can be found in **Table 1**.

The final MM-estimate $\hat{\boldsymbol{\theta}}_2$ inherits the high breakdown point of initial estimate because of re-descending psi-function ψ_{c_2} and the fixed scale $\hat{\sigma}_1$, but has a high efficiency.

Standard Errors

We report MM standard errors as returned by the lmRob function in R *robust* package. In particular, the three components of the covariance matrix, $\sigma^2 \tau \boldsymbol{C}_x^{-1}$, are estimated as follows.

The scale parameter σ is estimated by the initial scale estimate $\hat{\sigma}_1$, and the parameter τ is estimated by

$$\hat{\tau} = \frac{n}{n-p} ave_t \left\{ \left(\frac{\psi_{c_2}\left(r_t/\hat{\sigma}_1\right)}{ave_j \psi'_{c_2}\left(r_j/\hat{\sigma}_1\right)} \right)^2 \right\} \tag{A.2}$$

where $\psi_{c_2}(u) = \rho'_{c_2}(u)$, the r_t are residuals from the final MM fit and the factor $\frac{1}{n-p}$ is used to recapture the classical formula for LS for which $\psi(u) = u$. See Equation (5.33) in Maronna *et al.* [18].

Let $\boldsymbol{V}_{x^*} \equiv E\left(\boldsymbol{x}^* \boldsymbol{x}^{*\prime}\right) = \begin{bmatrix} 1 & \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x & E(\boldsymbol{xx'}) \end{bmatrix}$ with $\boldsymbol{\mu}_x \equiv E(\boldsymbol{x})$ denoting a vector of expected values of $\boldsymbol{x}$. The matrix block-inversion formula suggests that

$$\boldsymbol{V}_{x^*}^{-1} = \begin{bmatrix} 1 + \boldsymbol{\mu}'_x \boldsymbol{C}_x^{-1} \boldsymbol{\mu}_x & -\boldsymbol{\mu}'_x \boldsymbol{C}_x^{-1} \\ -\boldsymbol{C}_x^{-1} \boldsymbol{\mu}_x & \boldsymbol{C}_x^{-1} \end{bmatrix} \tag{A.3}$$

where $\boldsymbol{C}_x \equiv Var(\boldsymbol{x})$ is a covariance matrix of $\boldsymbol{x}$. Yohai *et al.* [26] proposed the following robust estimate of $\boldsymbol{V}_{x^*}$:

$$\hat{\boldsymbol{V}}_{x^*} = \frac{ave_t \left\{ \boldsymbol{x}_t^* \boldsymbol{x}_t^{*\prime} w_t \right\}}{ave_t w_t} \tag{A.4}$$

The robust weights w_t are needed to down-weight the influence of high-leverage $\boldsymbol{x}_t$ outliers when estimating the covariance matrix $\boldsymbol{V}_x$. lmRob uses weights computed from the initial S-estimate residuals and final MM estimate psi-function, *i.e.* $w_t = \frac{\psi_{c_2}\left(r_t^S/\hat{\sigma}\right)}{r_t^S/\hat{\sigma}}$. $\hat{\boldsymbol{C}}_x^{-1}$ is the corresponding subset of the $\hat{\boldsymbol{V}}_{x^*}^{-1}$. The $\hat{\boldsymbol{V}}_{x^*}$ is a consistent estimator of $\boldsymbol{V}_{x^*}$. Thus, by Slutsky theorem, $\hat{\boldsymbol{V}}_{x^*}^{-1}$ is a consistent estimator of $\boldsymbol{V}_{x^*}^{-1}$ and, consequently, $\hat{\boldsymbol{C}}_x^{-1}$ is a consistent estimator of $\boldsymbol{C}_x^{-1}$.

The test statistic T standard errors are obtained by multiplying the MM beta standard errors returned by lmRob by $\sqrt{1-EFF}$ where *EFF* is the normal distribution efficiency of the MM estimator.

Appendix C: Breakdown Point and Bias

The *breakdown point* (*BP*) of an estimate is defined as the smallest fraction of contamination that can cause the estimator to take on values arbitrarily far from its value for outlier free data. For example, moving a single data value to $\pm\infty$

cause the sample mean to move to ±∞, *i.e.*, BP = 0 for the sample mean. On the other hand, the sample median tolerates up to 50% of arbitrarily large outliers before it can be moved to ±∞, and, therefore BP =.5 for the sample median. See Hampel [32] for the introduction of the concept of breakdown point in robust statistics.

Let $\boldsymbol{\theta}_\infty$ be the asymptotic value of an estimator $\hat{\boldsymbol{\theta}}$, *i.e.*, $\hat{\boldsymbol{\theta}} \rightarrow_p \boldsymbol{\theta}_\infty$ (where $\rightarrow_p$ denotes convergence in probability), and $\boldsymbol{\theta}$ be the true parameter value. The *asymptotic bias* $B(\hat{\boldsymbol{\theta}})$ of a multivariate estimator $\hat{\boldsymbol{\theta}}$ may be usefully defined as

$$B(\hat{\boldsymbol{\theta}}) = \sqrt{(\boldsymbol{\theta}_\infty - \boldsymbol{\theta})' V_{x^*}(\boldsymbol{\theta}_\infty - \boldsymbol{\theta})}.$$ See Sections 3.3, 3.4, 3.6 and 6.7 in Maronna *et al.* [18].