Earnings to Price Analysis with mOpt versus Bisquare Robust Regression

Richard Douglas Martin[1], John B. Guerard[2], Daniel Z. Xia[3]

[1]Department of Applied Mathematics and Statistics, University of Washington, Seattle, USA
[2]Partner, Pacific Spirit Investments, Bluffton, USA
[3]Independent Researcher, Lone Tree, USA
Email: martinrd3d@gmail.com

Abstract

Recently Martin, Guerard, and Xia [1] used a new optimal bias robust regression estimator, called the mOpt estimator, in Fama-MacBeth cross-section regressions to study the statistical significance of the earnings-to-price (EP) and book-tp-price (BP) factors, among others. An earlier study by Markowitz *et al.* [2], and a number of studies referenced therein, used an alternative well-known Tukey Bisquare robust regression estimator. This begs the question of how the Bisquare estimator fares relative to the mOpt robust regression with regard to determining the statistical significance of the EP and BP factors. Here we show that the Bisquare robust regression estimator performs almost as well as mOpt with regard to the size of their significant t-statistics.

Keywords

Robust Regression, Cross-Section Factor Models, EP Factor, Outliers, Bias

1. The Fama-MacBeth Method and the Data

A wide variety of robust regression M-estimators have existed for some time, and are availiable in open source R packages such as RobStatTM and robustbase, and in commercial statistical software products such as Stata and SAS. In fact SAS offers 10 regression M-estimator weight function variants for its ROBUSTREG Procedure, with the following names: Andrews, Bisquare, Cauchy, Fair, Hampel, Huber, Logistic, Median, Talworth, Welsch. While SAS contains the Bisquare robust regression estimator studied herein, it does not contain the *mOpt* robust regression estimator studied herein, which however is contained in the RobStatTM package. The relatively new mOpt estimator is based on the theoretical results for an *Opt* optimal bias robust regression estimator discovered by Yohai

and Zamar [3], and modified by Konis and Martin [4]. The first, and only, currently published papers studying the use of the mOpt robust regression estimators in quantitative finance are those of Martin and Xia [5] and Martin *et al.* [6].

Recently Martin, Guerard, and Xia [1] used robust mOpt and least squares (LS) Fama-MacBeth cross-section factor model regressions to evaluate the statistical significance of EP and BP factors.[1] For at times $t = 1, 2, \cdots, T$ the factor model for EP and BP has the form

$$r_{i,t} = \theta_{0,t} + \sum_{k=1}^{2} b_{ik,t} \theta_{k,t} + s \cdot \varepsilon_{i,t}, i = 1, 2, \cdots, N_t \qquad (1)$$

where the $b_{ik,t}$ are the known values of *lagged factor exposures*, $\theta_{0,t}$ is the intercept, and $\theta_{1,t}, \theta_{2,t}$ are the regression slopes.[2] The parameter $s > 0$ is an unknown error term scale factor, which is needed for the mOpt robust regression method. The steps of the Fama-MacBeth method are: 1) Compute the intercept and regression slopes estimates $\hat{\theta}_{0,t}, \hat{\theta}_{1,t}, \hat{\theta}_{2,t}$ for each $t = 1, 2, \cdots, T$; 2) Use the time series of these slope estimates to compute heteroscedasticity and autocorrelation corrected (HAC) t-statistics for the intercept and slopes. The standard Fama-MacBeth method is based on LS regressions, and here we extend it to a *robust Fama-MacBeth* method based on the mOpt robust regression estimates described in Section 2.

We use the same stock data as in Martin, Guerard, and Xia [1], namely stock returns for the CRSP universe, and for emulated Russell R3000, R2000, and R1000 index universes, for the two time-periods 1980-2007, and 2008-2020. The average number of stocks in the cross-sections of those universes for 2008-2020 are 6961, 2769, 1833 and 967. The CRSP® data is from the Center for Research in Security Prices, LLC database, and the R3000, R2000 and R1000 stock universes were emulated from the CRSP database by matching the Russell indexes stock CUSIP's with those of the CRSP® database.

The EP and BP factor exposures are computed using balance sheet data from the Compustat database. For each calendar year, we use balance sheet data from the fiscal year ending in calendar year $t − 1$ for estimation starting in June of year t until May of year $t + 1$, predicting returns from July of year t until June of year $t + 1$. The EP value is the earnings-per-share (EPS) divided by the stock price at the end of each month, where EPS is firm's reported net income (NI) from its income statement, divided by the number of its common stocks outstanding. The BP value is the book value per share divided by the price per share, where the book value is the common stockholder equity (SEQ).

2. The mOpt and Bisquare Robust Regressions

The general form of a robust regression estimator for the regression model given

---

[1]Fama-MacBeth cross-section regressions are commonly used in *empirical asset pricing research*, and for an overview of this research area, see Bali, Engle, and Murray [7].

[2]Lagged factor exposures $b_{ik,t}$ are the exposure values measured at time $t − 1$. Since most of the factor exposures in this study are dimensionless ratios, and our main interest is in factor significance, we do not bother with the common practice of standardizing the exposures to having cross-section sample mean and standard deviations of 0 and 1, respectively.

by Equation (1) is as follows. Let $\boldsymbol{b}_{i,t} = \left(b_{i1,t}, b_{i2,t}\right)'$ be the vector of lagged exposures for the i-th stock at time t, and let $\tilde{\boldsymbol{b}}_{i,t} = \left(1, \boldsymbol{b}_{i,t}'\right)'$ be the lagged exposures vector augmented to include the intercept. Then with $\boldsymbol{\theta}_t = \left(\theta_{0,t}, \theta_{1,t}, \theta_{2,t}\right)'$, the 2-factor EP and BP model has the form:

$$r_{i,t} = \tilde{\boldsymbol{b}}_{i,t}'\boldsymbol{\theta}_t + s \cdot \varepsilon_{i,t}. \tag{2}$$

Dropping the time subscript t in the above model for notation convenience, the *regression residuals* are defined as

$$\hat{\varepsilon}_i\left(\boldsymbol{\theta}\right) = r_i - \tilde{\boldsymbol{b}}_i'\boldsymbol{\theta}, i = 1, 2, \cdots, N \tag{3}$$

where $\boldsymbol{\theta}$ is variable. It is shown in Martin *et al.* [6] that an mOpt robust regression M-estimator $\hat{\boldsymbol{\theta}}_t$ of $\boldsymbol{\theta}$ is a solution of the non-linear weighted least squares (WLS) estimating equation

$$\sum_{i=1}^{N} \mathrm{w}_{\mathrm{mOpt}}\left(\frac{\hat{\varepsilon}_i\left(\hat{\boldsymbol{\theta}}\right)}{\hat{s}}\right) \cdot \tilde{\boldsymbol{b}}_i\left(r_i - \tilde{\boldsymbol{b}}_i'\hat{\boldsymbol{\theta}}\right) = \boldsymbol{0} \tag{4}$$

where $\hat{s}$ is a robust residuals scale estimate computed prior to solving the above equation, and the weight function $\mathrm{w}_{\mathrm{mOpt}}\left(x\right)$ is determined by the derivative of the mOpt M-estimator loss function.[3] The formula for the mOpt weight function is

$$\mathrm{w}_{\mathrm{mOpt}}\left(x\right) = \begin{cases} 1 & |x| \le 1 \\ \dfrac{\phi(1)}{\phi(1) - 0.0132}\left(1 - \mathrm{SGN}(x)\dfrac{0.0132}{x\phi(x)}\right)\mathrm{U}\left(3.00 - |x|\right) & |x| > 1 \end{cases} \tag{5}$$

where $\mathrm{SGN}(x)$ is the "sign" function whose value is +1 for $x > 0$, −1 for $x < 0$, and 0 for $x = 0$, $\mathrm{U}(x)$ is the unit step function whose value is 1 for $x \ge 0$ and 0 for $x < 0$, and $\phi(x)$ is the standard normal probability density function. For details concerning the mOpt estimator weight function, see Konis and Martin [4].

The Equation (4) is solved for $\hat{\boldsymbol{\theta}}$ using an iterated weighted least square (IWLS) algorithm is briefly described in Equation (16) of Martin *et al.* [6], and described more completely in Martin, Guerard, and Xia [1]. Furthermore, proof of convergence of the IWLS algorithm for weight functions which are non-increasing in $|x|$ is provided in Section 9.1 of Maronna *et al.* [8].

The mOpt regression estimator is optimal in the sense of minimizing the maximum bias of $\hat{\boldsymbol{\theta}}$ due to joint factor exposure and return outliers. Details concerning this are provided in the Section "Efficient bias robustness of the

---

[3]A regression M-estimator $\hat{\boldsymbol{\theta}}_M$ minimizes the function $\sum_{i=1}^{N} \rho\left(\dfrac{\hat{\varepsilon}_i\left(\boldsymbol{\theta}\right)}{\hat{s}}\right)$, where loss function $\rho(x)$ is symmetric and non-decreasing with $\rho(0) = 0$. Differentiation of the summation shows that $\hat{\boldsymbol{\theta}}_M$ is a solution of the estimating equation $\sum_{i=1}^{N} \tilde{\boldsymbol{b}}_i \psi\left(\dfrac{\hat{\varepsilon}_i\left(\hat{\boldsymbol{\theta}}_M\right)}{\hat{s}}\right) = \boldsymbol{0}$, where

$\psi(x) = \rho'(x) = \dfrac{\mathrm{d}}{\mathrm{d}x}\rho(x)$. The weight function is then obtained from $\psi(x)$ as $\mathrm{w}(x) = \psi(x)/x$.

mOpt regression estimator" in Martin and Xia [5].

However, the robust regression M-estimator used in Markowitz *et al.* [2], and a number of references therein by John Guerard and co-authors, used the Tukey *Bisquare weight function*, whose formula is:

$$w_{\text{Bisquare}}(x) = \begin{cases} \left(1 - \left(\dfrac{x}{4.68} \right)^2 \right)^2 & |x| \le 4.68 \\ 0 & |x| > 4.68 \end{cases} \tag{6}$$

The above mOpt and Bisquare weight functions have been tuned, by the choice of their constants, to have 95% efficiency for the case where the ε_t in Equation (1) have a normal distribution. The shapes of $w_{\text{mOpt}}(x)$ and $w_{\text{Bisq}}(x)$ are displayed in **Figure 1**.

The mOpt weight function gives a weight of 1 to all robustly scaled residuals $\hat{\varepsilon}_t(\hat{\boldsymbol{\theta}}_t) = (r_i - \tilde{\boldsymbol{b}}_i'\hat{\boldsymbol{\theta}}_t)/\hat{s}$ that are less than 1 in magnitude, and smoothly transitions to a weight of 0 for robustly scaled residuals whose absolute value is greater than 3.00. Data returns and lagged exposures vector pairs $(r_t, \boldsymbol{b}_i)$ which result in 0 weights are said to be *rejected*. For normally distributed data and true parameter values, the probability that such a pair is rejected is a tiny 0.27%, and the estimator is virtually equivalent to the LS estimator.

The differences in the shapes of the mOpt and Bisquare weight functions in **Figure 1** suggest that the Bisquare robust regression will be sub-optimal, relative to the mOpt regression, in controlling bias due to outliers, because: The Bisquare weight function down-weights robustly scaled regression residuals $\hat{\varepsilon}_i(\boldsymbol{\theta})/\hat{s}$ *more* than the mOpt weight function for absolute values of the robustly scaled regression residuals greater than 0 and less than 3.0, and thereby are not considered to be outliers, and down-weights robustly scaled prediction residuals *less* than mOpt for robustly scaled prediction residuals with absolute values greater than 3.0, which are thereby considered to be outliers which are rejected by the mOpt regression.
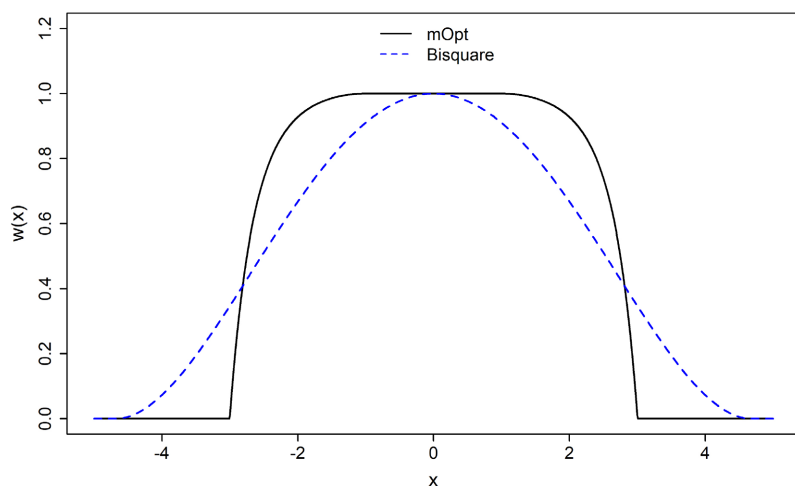
Figure 1. mOpt and Bisquare 95% normal distribution efficiency weight functions.

3. EP and BP Factor Significance with mOpt versus Bisquare

The study results reported in this Section are motivated by the recent results of Martin, Guerard, and Xia [1], who show that the earnings-to-price (EP) factor is not a significant factor for explaining the cross-section of returns when least-squares Fama-MacBeth regressions are used, but EP is highly significant when mOpt robust regressions are used. The EP factor was also shown to be an important factor in multi-factor models fit with Bisquare robust regressions, for the purpose of constructing mean variance optimal (MVO) portfolios in Markowitz *et al.* [2], and in papers co-authored by Guerard referenced therein. It has remained an open question of how well the Bisquare regressions perform relative to the optimal bias robust mOpt regressions, for the pupose of studying the significance of the EP factor.

In order to check on the extent to which the Bisquare regression is sub-optimal relative to the mOpt regression, we computed the HAC t-statistics (Tstats) for the Bisquare and mOpt regression for the CRSP*, R3000, R2000, R1000 universes and the time periods 1980-2007 and 2008-2020, and the results are displayed in Table 1. We consider Tstats with absolute values greater than 3.0 to be *significant*, and consider those with values at least 2.0 but less than 3.0 to be *weakly significant*, Thus in Table 1 we use green highlight for significant Tstats, and use yellow for weakly significant Tstats. The results show very clearly that EP is a significant using both mOpt and Bisquare regressions for the CRSP*, R3000 and R2000 universes for both time periods, but is only weakly significant for the R1000 during the first time period. Furthermore, for the mOpt and Bisquare regressions, BP is never significant, except for the CRSP universe for the second time period, where it is highly significant as a "negative value" factor. The CRSP data universe is curiously unique in this regard after 2007.

Table 1. Comparison of mOpt and Bisquare HAC Tstats for EP and BP multiple regressions for CRSP®, R3000, R2000, R1000 universes and the time periods 1980-2007 and 2008-2020.

Universe	mOpt EP	Bisquare EP	mOpt BP	Bisquare BP
1980-2007				
CRSP	**6.09**	**6.28**	2.03	1.86
R3000	**6.86**	**6.23**	0.46	0.47
R2000	**7.1**	**6.49**	1.24	1.12
R1000	2.67	2.61	0.21	0.16
2008-2020				
CRSP	**6.07**	**5.98**	**−4.75**	**−4.78**
R3000	**6.07**	**5.82**	−1.16	−0.97
R2000	**6.46**	**6.07**	−0.52	−0.46
R1000	1.12	0.94	−1.1	−0.95

The mOpt Tstat is larger than that of the Bisquare for all but 3 of the 16 pairwise comparisons of the two Tstats, and for 2 of those 3 it is a tiny difference in the second digit. The mean relative difference between the mOpt and Bisquare Tstats is 9.2%, and minimum and maximum relative differences of −3.2% and 31.2%. However, there is only one case where the significance outcomes are different, namely for BP using the CRSP˚ universe in the first time period where the mOpt Tstat is weakly significant but the Bisquare is not at all significant. Thus one can conclude that most, if not all, of the conclusions based on applying the Bisquare robust regression to minimum variance portfolio optimization in Markowitz *et al.* [2], in papers co-authored by Guerard therein, would be changed very little, if at all, by using the theory based mOpt regressions. However, it is preferable to use an estimator such as the mOpt estimator, for which there exists solid theoretical support.

4. Concluding Comments

We have shown that the Bisquare regressions performance in terms of the size of HAC t-statistics is not as large as those of the mOpt robust regression, but they are not much smaller. These empirical results are consistent with the theoretical optimal bias robustness of the mOpt regression estimator, which we strongly recommend for quantitative finance research and applications which involve time series and cross-section factor models. For an introduction to robust statistics for portfolio construction and analysis we recommend Martin *et al.* [6], and for robust time series factor models see Martin and Xia [5].

One cannot stress too strongly the usefulness of mOpt robust regression as a diagnostic tool for checking whether or not a least squares regression has been unknowingly influenced by outliers. The software for computing mOpt robust regressions is available in the form of the lmrobdetMM function in the R open source R package RobStatTM, downloadable at https://cran.r-project.org/web/packages/RobStatTM.

Conflicts of Interest

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

References

[1] Martin, R.D., Guerard, J.B., and Xia, D.Z. (2024) Resurrecting Earnings-to-Price with Machine Learning Robust Control for Outliers. https://ssrn.com/abstract=4746580

[2] Markowitz, H., *et al.* (2021) Financial Anomalies in Portfolio Construction and Management. *The Journal of Portfolio Management*, **47**, 51-64. https://doi.org/10.3905/jpm.2021.1.242

[3] Yohai, V.J. and Zamar, R.H. (1997) Optimally Local Robust M-Estimates of Regression. *Journal of Statistical Planning and Inference*, **64**, 309-323. https://doi.org/10.1016/S0378-3758(97)00040-2

[4] Konis, K. and Martin, R.D. (2021) Optimal Bias Robust Psi and Rho Revisited. https://ssrn.com/abstract=3902862

[5] Martin, R.D. and Xia, D.Z. (2022) Efficient Bias Robust Regression for Time Series Factor Models. *Journal of Asset Management*, **23**, 215-234. https://link.springer.com/content/pdf/10.1057/s41260-%20022-00258-0.pdf

[6] Martin, R.D., *et al.* (2023) Robust Statistics for Portfolio Construction and Analysis. *The Journal of Portfolio Management*, **49**, 105-139. https://doi.org/10.3905/jpm.2023.1.527

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

[8] Maronna, R.A., *et al.* (2019) Robust Statistics: Theory and Methods (with R). 2nd Edition, John Wiley & Sons, New York. https://doi.org/10.1002/9781119214656