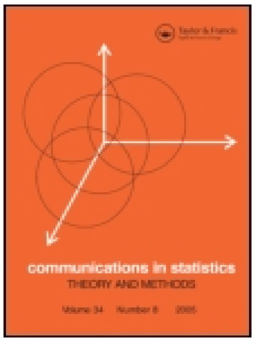

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

The method of moments ratio estimator for the tail shape parameter

Jon Danielsson , Dennis W. Jansen & Casper G. De vries

To cite this article: Jon Danielsson , Dennis W. Jansen & Casper G. De vries (1996) The method of moments ratio estimator for the tail shape parameter, Communications in Statistics - Theory and Methods, 25:4, 711-720, DOI: 10.1080/03610929608831727

To link to this article: https://doi.org/10.1080/03610929608831727

Published online: 27 Jun 2007.

COMMUN. STATIST.—THEORY METH., 25(4), 711-720 (1996)

THE METHOD OF MOMENTS RATIO ESTIMATOR FOR THE TAIL SHAPE PARAMETER

Jon Danielsson

University of Iceland
Department of Economics, 101 Reykjavik, Iceland

Dennis W. Jansen

Texas A&M University
Department of Economics, College Station, Texas 77843-4228, USA

Casper G. de Vries

Tinbergen Institute and Erasmus University Rotterdam
Oostmaaslaan 950, 3063 DM Rotterdam, The Netherlands

Key Words and Phrases: Hill estimator, method of moments, bias

ABSTRACT

The so-called Hill estimator for the shape parameter of the tail distribution is known to be downwardly biased. The Hill estimator is a moment estimator, based on the first conditional moment of the highest logarithmically transformed data. We propose a new estimator for the tail index based on the ratio of the second to the first conditional moment. This estimator has a smaller bias than the Hill estimator. We provide simulation results that demonstrate a sizable reduction in bias when α is large, while the MSE is moderated as well. The new estimator is applied to stock return data in order to resolve a long standing issue in economics.

1. INTRODUCTION

It is well known from the work of Hall (1982) that the Hill (1975) estimator can be severely biased. In order to investigate this bias, we look at the second order expansions of regularly varying distribution functions. For β not zero, the expansion:

$$F(x) = 1 - ax^{-\alpha}[1 + bx^{-\beta} + o(x^{-\beta})], \tag{1.1}$$

holds as $x \to \infty$, and where $\alpha, \beta > 0$, $a > 0$, and $b \in \mathbb{R}$. Corresponding to this is the density function, given by:

$$f(x) = \alpha a x^{-\alpha-1} + (\alpha + \beta) abx^{-(\alpha+\beta)-1} + o(x^{-(\alpha+\beta)-1}). \tag{1.2}$$

Here we consider the Hill estimator of the tail shape parameter α, given by

$$\gamma = 1/\alpha = (1/m)\,\Sigma \ln(X_{[n-i]}/X_{[n-m]}), \tag{1.3}$$

where the $X_{[n-1]}$ are the descending order statistics, $X_{[n]} > X_{[n-1]} > X_{[n-2]} > \ldots$ from a sample $X_1, X_2, \ldots, X_n$, and where $n > m > 1$. For some distributions, e.g. the Pareto distribution (in which $b = 0$) the Hill estimator is unbiased and consistent, but for $b \neq 0$ the Hill estimator is biased and may be inconsistent, depending on the rate at which $m \to \infty$ as $n \to \infty$. In fact Hall (1982) showed that the asymptotic MSE is minimized if the bias squared and variance vanish at the same rate. Moreover, Hall and Welsh (1984) show that this rate is the best attainable rate. This implies that other estimators may differ from the Hill estimator only with respect to the constant factor in the asymptotic MSE.

What is the bias in the Hill estimator? Let s be some large threshold, and suppose that all order statistics which exceed s are used to calculate γ. Using (1.2) and (1.3), the mean of the Hill estimator can then be calculated as:

$$E(\gamma \mid s) = (1/\alpha) - \beta bs^{-\beta}/[\alpha(\alpha + \beta) + \alpha(\alpha + \beta)bs^{-\beta}] + o(s^{-\beta}), \tag{1.4}$$

see Goldie and Smith (1987). From (1.4) the bias is easily isolated. Simulation results reported below suggest that the bias in the Hill estimator can be considerable, especially for values of α larger than 2. However, since for financial data there are numerous reported estimates of α greater than 2, but usually less than 4 or 5, it is of interest to examine the properties of the Hill estimator for distributions with α in this range, and to see if an alternative estimator might outperform the Hill estimator in these situations.

2. AN ALTERNATIVE TO THE HILL ESTIMATOR

As an alternative for the Hill estimator, we first considered a moment estimator, which conditions on a high threshold value s, and takes into account the second order term that gives the bias in the Hill estimator. Assume the second order expansion in (1.1) is exact. Then the conditional moments of the log transformed data are

$$E\left[\log(X_i/s)^k \mid X_i \geq s\right] = \frac{k!}{1+bs^{-\beta}}\left[\frac{1}{\alpha^k} + \frac{bs^{-\beta}}{(\alpha+\beta)^k}\right] \tag{2.1}$$

and the empirical conditional moments may be calculated from:

$$\tilde{m}^{(k)} = \frac{1}{m}\sum_{i=1}^{m}(\log X_i/s)^k \mid X_i > s, \text{ and } m = \#X_i > s. \tag{2.2}$$

Equating (2.2) to (2.1) for $k = 1,2$ and 3, one can solve out for α, β, and b in terms of $\tilde{m}^{(1)}, \tilde{m}^{(2)}$ and $\tilde{m}^{(3)}$. Unfortunately, unpublished simulations showed that this exact moment estimator is numerically quite unstable even for large data sets (such as $n = 10^5$).

The formula (2.1) for the conditional moments, however, suggests an alternative moment estimator. Note that in equation (2.1) the second term has the same sign for all k. Thus the bias term in (1.4) is directly related to the second term in (2.1) for $k = 1$, but also for $k > 1$. In particular for $k = 1,2$ we have

$$E[\log(X_i/s) \mid X_i \geq s] = \frac{1}{1+bs^{-\beta}}\left[\frac{1}{\alpha} + \frac{bs^{-\beta}}{(\alpha+\beta)}\right] \tag{2.3}$$

and the second conditional moment is given by

$$E[\log(X_i/s)^2 \mid X_i \geq s] = \frac{2}{1+bs^{-\beta}}\left[\frac{1}{\alpha^2} + \frac{bs^{-\beta}}{(\alpha+\beta)^2}\right]. \tag{2.4}$$

Taking conditional expectations in (1.3) it follows that the first conditional moment (2.3) is just the conditional expectation of the Hill estimator, and that the second term in (2.4) goes in the same direction as the bias term in (2.3). Also, in the expression for the second conditional moment, the first term is $1/\alpha$ squared. The idea is then to use the ratio of the two empirical moments to estimate $1/\alpha$. The moments ratio estimator (*MRE*) is defined as

$$MRE_j \equiv \frac{1}{(1+j)} \frac{\tilde{m}^{(j+1)}}{\tilde{m}^{(j)}}. \tag{2.5}$$

For example, MRE_1 will be the ratio of the second conditional moment to the first. Taking $s \sim cn^{1/(2\beta+\alpha)}$ as $n \to \infty$, see Goldie and Smith (1987), one finds that

$$MRE_j \to \frac{1}{\alpha} - \frac{b}{s^\beta} \frac{\beta\alpha^{j-1}}{(\alpha+\beta)^{j+1}} \text{ as } n \to \infty. \tag{2.6}$$

From (2.6) the asymptotic bias for $j = 1$ is as follows

$$-\beta bs^{-\beta}/(\alpha+\beta)^2. \tag{2.7}$$

Comparing the bias part in (1.4) of the Hill estimator to the bias (2.7) of the MRE_1, it is straightforward to show that the MRE_1 has a lower asymptotic squared bias in comparison with the Hill estimator when evaluated at the same threshold. But the convergence rates are still the same, see Hall and Welsh (1984). In fact Danielsson et al (1994) and De Haan and Peng (1994) show that if $\alpha > \beta$, the MRE_1 has lower bias and MSE than the Hill estimator when both are evaluated at their respective MSE minimizing thresholds.

3. MONTE CARLO RESULTS

To investigate the finite sample properties of the estimators (2.5), we set up a Monte Carlo experiment. We simulated with a stable distribution for which $\alpha = 0.5$ and with student-t distributions with 1, 2, 3, 4, and 5 degrees of freedom (and recall α is equal to these degrees of freedom for the student-t distributions). The simulation involved 50,000 observations for each run, with 500 replications.

We varied the number of order statistics used to calculate our estimates of α from .1% of the sample to 15% of the sample. In consideration of space, we limit ourselves in the tables below to reporting the results for .5%, 1%, 2%, 3%, 5%, and 10% of the sample. For each replication we computed the bias and the MSE. We also ran these same simulations using 200,000 observations for each run with essentially the same results. Simulations with 6,000 observations for the Hill estimator only are reported in Jansen and de Vries (1991).

In Table 1 we report the bias with respect to α (and not its inverse γ) as decimal places. Starting the discussion with the Hill statistic, there are two things to notice. First, the bias can be substantial, especially for the higher α values, and is negative. The downward bias follows directly from the fact that the parameter b in the expansion of the student-t density is negative ($b = -\alpha^2(\alpha+1)/[2(\alpha+2)]$). The second point to notice is that the MSE is a U-shaped function of the number of order statistics. This is a direct implication of theorem 2 in Hall (1982). Turning towards the MRE estimators, the same observations hold. But there are some interesting differences as well.

In order to compare the different estimators, we focus on the number of order statistics that minimize the respective MSE's. For example, if $\alpha = 2$, the Hill statistic has minimal MSE at 2% of the data, while MRE_1 has minimal MSE at 3% of the data. At both of these thresholds the MRE_1 has the lower bias squared, cf. the discussion following equation (2.7) above. But it is also the case that when both estimators are evaluated at their optimal threshold (in the MSE sense), the MRE_1 produces the lower bias (and a lower MSE as well for values of α above 1). This is a general pattern. Also note that MRE_2 and MRE_3 perform worse than MRE_1 in the MSE sense, but can perform better (at the same threshold level) in reducing the size of the bias. This issue is further discussed in Danielsson et al (1994).

Table 2 gives some information on the shape of the sample distribution of the Hill and MRE_1 estimators when the underlying population is student-t with 2 degrees of freedom. The Hill estimator was calculated using a 2% threshold, while MRE_1 was calculated using a 3% threshold. The bias is evident, as is the lack of skewness or kurtosis. The quantiles evidence the symmetry of the distribution.

Finally, in Table 3 we provide a small example applying these estimators to 2004 monthly observations on the U.S. stock market index compiled by Schwert

TABLE 1. Simulation Results

No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	Stable Distribution ($\alpha = 0.5$)							
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	0.0007	0.0009	0.0057	0.0019	0.0179	0.0048	0.0434	0.0104
1.0%	0.0000	0.0005	0.0023	0.0009	0.0092	0.0025	0.0253	0.0057
2.0%	0.0000	0.0002	0.0008	0.0004	0.0045	0.0013	0.0142	0.0032
3.0%	0.0000	0.0001	0.0003	0.0003	0.0028	0.0008	0.0099	0.0023
5.0%	0.0000	0.0001	0.0000	0.0002	0.0015	0.0005	0.0061	0.0014
10.0%	0.0000	0.0000	0.0000	0.0001	0.0005	0.0002	0.0030	0.0008
	Student t Distribution ($\alpha = 1$)							
No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	0.0019	0.0040	0.0099	0.0086	0.0355	0.0231	0.0893	0.0487
1.0%	-0.0002	0.0020	0.0035	0.0041	0.0172	0.0119	0.0507	0.0275
2.0%	0.0001	0.0010	0.0012	0.0021	0.0077	0.0061	0.0271	0.0155
3.0%	-0.0010	0.0006	0.0003	0.0014	0.0045	0.0041	0.0182	0.0110
5.0%	-0.0050	0.0004	-0.0010	0.0008	0.0019	0.0024	0.0106	0.0070
10.0%	-0.0210	0.0006	-0.0070	0.0004	-0.0020	0.0012	0.0040	0.0037
	Student t Distribution ($\alpha = 2$)							
No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	-0.0080	0.0166	0.0100	0.0300	0.060	0.0721	0.1625	0.1556
1.0%	-0.0259	0.0084	-0.0060	0.0150	0.0245	0.0372	0.0894	0.0849
2.0%	-0.0576	0.0068	-0.0257	0.0079	0.0005	0.0189	0.0427	0.0460
3.0%	-0.0881	0.0100	-0.0427	0.0065	-0.0130	0.0127	0.0225	0.0316
5.0%	-0.1484	0.0232	-0.0756	0.0083	-0.0341	0.0084	0.0000	0.0193
10.0%	-0.3013	0.0912	-0.1614	0.0270	-0.0840	0.0101	-0.0366	0.0102

(continued)

(1989), and augmented with data from the CRSP dataset. The sample runs from January 1826 through December 1992. The first column reports the number of lowest order statistics that are used in computing the estimates. We look only at the lower tail of the distribution of returns, defined as the differences of natural logs of the stock index. Prior work by Jansen and de Vries (1991) and Loretan

TABLE 1. (Continued) Simulation Results

	Student t Distribution ($\alpha = 3$)							
No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	-0.1138	0.0477	-0.0392	0.0635	0.0590	0.1432	0.2233	0.3225
1.0%	-0.1900	0.0515	-0.1024	0.0416	-0.0188	0.0744	0.0971	0.1663
2.0%	-0.2947	0.0937	-.1797	0.0464	-0.0909	0.0452	0.0029	0.0860
3.0%	-0.3883	0.1544	-0.2413	0.0667	-0.1385	0.0427	-0.0469	0.0603
5.0%	-0.5491	0.3036	-0.3484	0.1255	-0.2152	0.0586	-0.1140	0.0461
10.0%	-0.8675	0.7534	-0.5754	0.3325	-0.3776	0.1468	-0.2384	0.0695
	Student t Distribution ($\alpha = 4$)							
No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	-0.3576	0.1851	-0.2208	0.1537	-0.0800	0.2566	0.1285	0.4988
1.0%	-0.5102	0.2859	-0.3423	0.1669	-0.2040	0.1670	-0.0452	0.2749
2.0%	-0.7213	0.5298	-0.5005	0.2710	-0.3371	0.1708	-0.1911	0.1786
3.0%	-0.8790	0.7779	-0.6188	0.3946	-0.4301	0.2188	-0.2772	0.1681
5.0%	-1.1307	1.2811	-0.8092	0.6602	-0.5778	0.3499	-0.4018	0.2088
10.0%	-1.5957	2.5472	-1.1777	1.3887	-0.8704	0.7623	-0.6388	0.4229
	Student t Distribution ($\alpha = 5$)							
No. of order stats	Hill		MRE_1		MRE_2		MRE_3	
	bias	MSE	bias	MSE	bias	MSE	bias	MSE
0.5%	-0.7362	0.6130	-0.5090	0.3708	-0.2847	0.3546	0.0025	0.5735
1.0%	-0.9753	0.9803	-0.7105	0.5573	-0.4845	0.3588	-0.2498	0.3501
2.0%	-1.2810	1.6534	-0.9567	0.9378	-0.7011	0.5451	-0.4752	0.3574
3.0%	-1.5048	2.2710	-1.1355	1.3019	-0.8509	0.7552	-0.6137	0.4552
5.0%	-1.8399	3.3885	-1.4100	1.9937	-1.0802	1.1813	-0.8152	0.7020
10.0%	-2.4142	5.8296	-1.9025	3.6214	-1.5033	2.2640	-1.1827	1.4095

and Phillips (1994) using the Hill estimator suggests that α lies between 2 and 5, and that point estimates are close to 3. For a long time economists have debated the finiteness of the second moment, and more recently of the fourth moment. Thus removing possible bias in the α-estimates is important. Our simulation results suggest using .5% of the sample for the Hill estimator when α is 3 and the distribution of returns is student-t. For the MRE_1 estimator our simulation suggests using order statistics comprising 1% of the sample, while for MRE_2 the simulation suggests using 3% of the sample, and 5% for MRE_3. Thus we

TABLE 2. Simulated Distribution of $\hat{\alpha}$ from Student t ($\alpha = 2$)

	Hill Estimator	MRE_1 Estimator
Mean	1.9386	1.9541
Variance	.0019	.0023
Skewness	-.0241	.1526
Kurtosis	-.0346	.3992
N	500	500
Quantiles		
Max	2.0674	2.1463
75%	1.9685	1.9866
Median	1.9396	1.9515
25%	1.9067	1.9221
Min	1.7843	1.8159

TABLE 3. Estimates of α for the Lower Tail of Stock Returns

2004 monthly observations on U.S. stock market index, 1826.01 - 1992.12.

Number of Order Statistics	Hill Estimator	MRE_1 Estimator	MRE_2 Estimator	MRE_3 Estimator
10	2.881	4.762	6.180	7.406
20	2.785	3.517	4.475	5.501
40	2.666	3.081	3.667	4.378
60	2.773	2.990	3.430	4.022
100	2.416	2.795	3.134	3.557

calculated the various estimators for 10, 20, 40, 60, and 100 order statistics. Our estimates of α are reported in Table 3. Notice that the Hill estimator is 2.881 using 10 order statistics, and only varies between 2.416 and 2.881 as we increase the number of order statistics from 10 to 100. For the MRE_1 estimator using 20 order statistics, our estimate of α is 3.517, and the estimate falls from 4.762 to 2.795 as we increase the order statistics from 10 to 100. For MRE_2 the estimate of α using 60 order statistics is 3.430, and this estimate falls from 6.180 to 3.134 as we increase the number of order statistics from 10 to 100. Finally, for MRE$_3$ the estimate of α is 3.557 using 100 order statistics, and the estimate also falls from 7.406 to 3.557 as we increase the number of order statistics used. Thus, based on prior information that α might be around 3, and using our simulation to choose the number of order statistics, we get fairly close estimates of α from the Hill estimator and the three moment ratio estimators. But the MRE's are somewhat above 3, while the Hill estimator is somewhat below 3. Note, however, it is still the case that the choice of the number of order statistics is important not just when using the Hill estimator but also in using the moment ratio estimators. It follows that the variance of the stock returns is finite but the fourth moment seems unbounded.

5. CONCLUSION

To conclude, it is known that the Hill (or first conditional moment) estimator can be severely biased. We have introduced a novel estimator based on the ratio of the second to the first conditional moment that lessens this bias. Simulations on the student-t model confirm this, and also show an improvement in terms of MSE. An application to economic data was provided. Future work will have to focus on the issue of choosing the optimal number of order statistics empirically, and providing a fuller comparison of the properties of the Hill estimator and the MRE's.

ACKNOWLEDGMENTS

The authors thank a referee for constructive comments. Danielsson and Jansen thank the Tinbergen Institute for its hospitality and Jansen thanks the Private Enterprise Research Center at Texas A & M for research support.

BIBLIOGRAPHY

Danielsson, J., D. Jansen and C. G. de Vries (1994). "The moments ratio approach to quantile and tail estimation," mimeo, Tinbergen Institute Rotterdam.

Goldie, Charles M. and Richard L. Smith (1987). "Slow variation with remainder: theory and applications," *Quarterly Journal of Mathematics*, 38, 45-71.

Hall, Peter, and A. H. Welsh, (1984). "Best attainable rates of convergence for estimates of parameters of regular variations," *Ann. of Statis.*, 12, 1079-1084.

Hall, Peter (1982). "On some simple estimates of an exponent of regular variation," *J. Roy. Statis. Soc.* B, 44, 37-42.

Haan, de, L. and L. Peng (1994). "Comparison of tail index estimators," mimeo, Economic Institute, Erasmus University, Rotterdam.

Hill, Bruce M.(1975). "A simple general approach to inference about the tail of a distribution," *Ann. of Statis.*, 3, 1163-1173.

Jansen, Dennis W. and Casper G. de Vries (1991). "On the frequency of large stock returns: Putting booms and busts into perspective," *The Review of Economics and Statistics*, 73, 18-24.

Loretan, Mico, and Peter C.B. Phillips (1994). "Testing the covariance stationarity of heavy-tailed time series," *Journal of Empirical Finance*, 1, 211-248.

Schwert, G. William (1989). "Business cycles, financial crises and stock volatility," *Carnegie- Rochester Conference Series on Public Policy*, 31, 83-126.