# Online Appendix

# Covariates hiding in the tails

Milian Bachem*

Erasmus University Rotterdam

Lerby M. Ergun†

Bank of Canada
London School of Economics

Casper G. de Vries‡

Erasmus University Rotterdam

Friday 23rd June, 2023

*bachem@ese.eur.nl, Burgemeester Oudlaan 50, Rotterdam, 3062 PA, The Netherlands.

†Corresponding author: lerbyergun@gmail.com, 234 Wellington Street West, Ottawa, ON, K1A 0G9, Canada.

‡cdevries@ese.eur.nl, Burgemeester Oudlaan 50, Rotterdam, 3062 PA, The Netherlands.

## Bias under second-order Hall expansion

Recall the so-called Hall expansion (Hall and Welsh, 1985):

$$F(x) = 1 - Cx^{-\alpha}[1 + Dx^{-\theta} + o(x^{-\theta})].$$

Here $\alpha > 0$, $C > 0$, $\theta > 0$ and $D$ is a real number. Here $C$ and $D$ are the first- and second-order scale parameters, where $\alpha$ and $\theta$ are the first- and second-order shape parameters. We give a short derivation of the (conditional) bias in case the distribution function adheres to the above expansion. If the distribution function satisfies the monotone density theorem (see Bingham et al., 1987), it is sufficiently smooth so that the derivative gives its density with tail expansion

$$f(x) = \alpha Cx^{-\alpha-1} + (\alpha + \theta)CDx^{-\alpha-\theta-1} + o(1).$$

The conditional expectation can now be found as follows:

$$\begin{aligned}&E[\ln \frac{Y}{u}|Y > u]\\ &\simeq \frac{1}{Cu^{-\alpha}[1 + Du^{-\theta}]} \int_u^\infty (\ln \frac{x}{u}) \left[\alpha Cx^{-\alpha-1} + (\alpha + \theta)CDx^{-\alpha-\theta-1}\right] dx,\end{aligned}$$

if we omit the terms that are of order small. Note that we can cancel the $C$ factor from the numerator and denominator. If we then apply the calculus result

$$\alpha \int_u^\infty (\ln \frac{s}{u})s^{-\alpha-1}ds = -(\ln \frac{s}{u})s^{-\alpha}|_u^\infty + \int_u^\infty s^{-\alpha-1}ds = \frac{1}{\alpha}u^{-\alpha}$$

to the two parts separately, we obtain

$$\begin{aligned}E[\ln \frac{Y}{u}|Y > u] &\simeq \frac{u^\alpha}{1 + Du^{-\theta}} \left[\frac{1}{\alpha}u^{-\alpha} + \frac{1}{\alpha + \theta}Du^{-\alpha-\theta}\right]\\ &= \frac{1}{\alpha} + (\frac{1}{\alpha + \theta} - \frac{1}{\alpha})\frac{Du^{-\theta}}{1 + Du^{-\theta}}\\ &\simeq \frac{1}{\alpha} - \frac{\theta}{\alpha(\alpha + \theta)}Du^{-\theta}\end{aligned}$$

as $1 + Du^{-\theta} \to 1$ for $u$ large.

## Simulations

To investigate the efficacy of the above methods of bias reduction, we conduct simulations. In the first model, we simulate a cross section by adding a constant (positive or negative) to a heavy-tailed innovations term. The second approach simulates from a linear factor model, where factor $g$ is multiplied with a coefficient $\gamma_j$ specific to entity $j$.

Thus, in the first model data are simulated from

$$Y_j = h + X_j, \tag{1}$$

where $h$ is constant and $j = 1, ..., 20000$ signifies the cross section. The $X_j$ are heavy-tailed innovations drawn from a Student-t distribution with 3 degrees of freedom.[1]

In the second model, data are simulated from

$$Y_j = \gamma_j g + X_j, \tag{2}$$

where $X_j \sim$ Student-t and where $\gamma_j \sim \mathcal{N}(1, 0.25)$.[2] We again take $j = 1, ..., 20000$.

Note from the main text the bias in the right tail:

$$E\left[\ln \frac{Y}{u} | Y > u\right] = \frac{1}{\alpha} - \frac{1}{\alpha + 1} h u^{-1} + o\left(u^{-1}\right). \tag{3}$$

[1]Note that in finite samples, bias will remain due to the fact that we simulate from a Student-t distribution; the Hill estimator is only unbiased when data are drawn from the Pareto distribution.

[2]To bring the simulations closer to our empirical application we also draw $\gamma_j$ from a subsample of estimated CAPM betas ($n = 13,535$) for US stocks obtained from the CRSP database. The distribution of the CAPM betas has a mean of 1.14 and a standard deviation of 0.72. The unreported results are very similar and are available on request.

and the bias in the left tail:

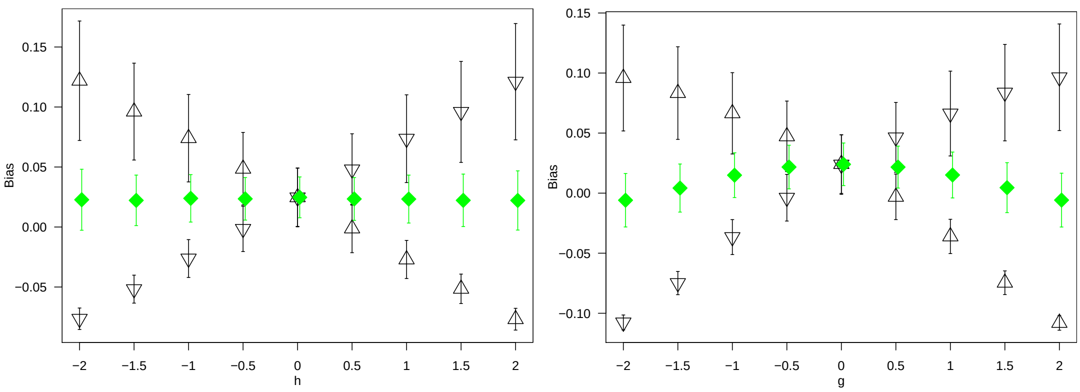
$$E\left[\ln\frac{Y}{u}|Y\leqslant -u\right]=\frac{1}{\alpha}+\frac{1}{\alpha+1}hu^{-1}+o\left(u^{-1}\right). \tag{4}$$

The above can be used to study the efficacy of the mirror method of bias reduction. To examine the efficacy of the shift method in the case of tail asymmetry, we simulate from a Student-t (2.5) and a Student-t (3.5) for the left and right side of the distribution, respectively. In the calculation of Hill estimates, we set the threshold $u$ at 4.54, which corresponds to the 99% quantile for a Student-t (3) distribution.

## Mirror method

Figure 1 presents cross-sectional Hill estimates on the simulated data for the left and the right tail with the inclusion of bias-corrected estimates using the mirror method. Panel (a) shows the Hill estimates of the model with the deterministic shift described in (1). The value of the bias in left (right) tail index estimates is signified by the downward (upward) pointing triangles. In the centre of the figure, where $h = 0$, the average of left and right tail estimates are close to identical. The slight positive bias in both estimates is due to the second-order term in the tail expansion of the Student-t (3) innovation terms. Furthermore, the pattern of the triangles confirms the linear relationship between the value of $h$ and the bias in Hill estimates as derived in (3) and (4). In the left tail an increase in $h$ produces an increased bias, as shown in (4). This opposite relationship can be observed for the right tail, as in (3). Since the Hill estimates change at the same (absolute) linear rate in the left and the right tail, bias correction using the mirror method produces estimates (green diamonds) that are close to being unbiased. Furthermore, the bars surrounding the triangles indicate that the variances of $\hat{\alpha}^Y_-$ and $\hat{\alpha}^Y_+$ differ. This is due to the fact that the bias shifts the estimates away from $\alpha$ and the variance is negatively correlated to the level of $\alpha$. The mirror method reduces the standard deviation of the estimates by $1/\sqrt(2)$ relative to the average of uncorrected left and right tail estimate standard deviation as shown in the asymptotic distribution for the mirror method in the main text. This is most clearly demonstrated by the difference in the size of the black and green bars at $h = 0$.

Figure 1: Bias correction for $1/\hat{\alpha}$ using the mirror method




(a) Constant model (b) Linear model

This figure presents the results of bias correction using the mirror method discussed in Section **??**. In panels (a) and (b), data are simulated from the models described in (1) and (2), respectively, where the $X_j$ are drawn from a Student-t (3). The x-axis gives different values of the deterministic constant $h$ in panel (a) and different values of factor $g$ in panel (b). The y-axis indicates the bias by subtracting 1/3 from the Hill estimates with threshold $u = 4.54$ (99% quantile of the Student-t (3)). The upward pointing black triangles (△) show the bias of uncorrected Hill estimates in the right tail, the downward pointing black triangles (▽) show the bias of uncorrected Hill estimates in the left tail and the green diamonds (♦) show the bias of corrected Hill estimates using the mirror method. The bars surrounding the mean estimates present the standard deviation of the estimates.

Panel (b) presents the Hill estimates of the linear model described in (2). For the model with entity specific shift, the relationship between the value of the bias and $g$ is no longer linear. In a finite sample, for intermediate values of $u$, the contribution of $\gamma_i$ is non-negligible. In the above simulations, $\gamma_i g \sim N(g\mu_\gamma; g^2\sigma_\gamma^2)$. For large $g$, large values of $\gamma_i$ contribute to the ranking of the observations in the tail. This adds a thin tailed, normally distributed component, to the otherwise heavy-tailed $Y_j$ observations. This biases the $1/\hat{\alpha}^-$ and $1/\hat{\alpha}^+$ downwards.[3] Consequently, the mirror method estimates are also revised downwards as $|g|$ becomes larger. However, the influence of the coefficients is moderate to small on the efficiency of the mirror method.

Using the relationship between (3) and (4) to remove the cross-sectional location

[3]Although the normal distribution is outside of the support of the Hill estimator, using normally distributed data for the Hill estimator produces values close to zero and therefore biases the estimates downwards.

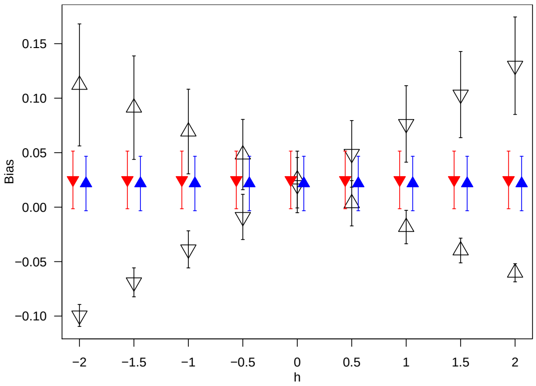
shift works under tail symmetry only. Under tail asymmetry there are three problems that surface. The obvious one is that the estimate neither reflects the left nor the right tail index. The more subtle effect is that $1/(1+\alpha)$ in (3) and (4) are different in the left and the right tail, i.e., $\alpha_-$ and $\alpha_+$ differ. This induces a different effect of $h$ in the left and right tail. Additionally, the optimal threshold $u$ may differ for the left and right tail. The optimal threshold $u$ varies inversely with the tail index $\alpha$, thus it may not be optimal to use the same threshold for both tails.

## Shift method

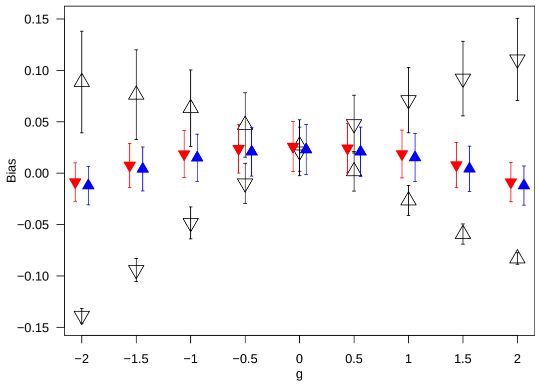
The cross-sectional Hill estimates for the case with asymmetric tails is presented in Figure 2. Panel (a) shows the Hill estimates for the model with a deterministic shift described in (1). The bias-corrected estimates using the shift method for the left (right) tail are indicated by the red downward (blue upward) pointing triangles. Bias correction using the shift method also produces estimates with near zero bias, despite the asymmetry between the bias in Hill estimates on the left and the right tail.

Panel (b) presents the Hill estimates of the linear model described in (2). Due to the inclusion of an entity specific shift term, the value of the bias has again become dependent on the values of the coefficients in the tail. By moving from a deterministic shift to an entity specific shift the efficacy of bias correction decreases somewhat. However, bias correction remains highly effective, producing tail estimates with far smaller bias than the uncorrected estimates.

Figure 2: Bias correction using the shift method




(a) Constant model




(b) Linear model

This figure presents the results of bias correction using the shift method discussed in Section **??**. In panels (a) and (b), data are simulated from the models described in (1) and (2), respectively. The $X_j$ innovations are drawn from a Student-t (2.5) and (3.5) for the left and right tail, respectively. The x-axis gives different values of the deterministic constant $h$ in panel (a) and different values of factor $g$ in panel (b). The y-axis indicates the bias by subtracting 1/2.5 and 1/3.5 from the Hill estimates in the left and right tail, respectively. The threshold is set at $u = 4.54$ (99% quantile of the Student-t (3)). The upward pointing black triangles (△) show the bias of uncorrected Hill estimates in the right tail, and the downward pointing black triangles (▽) show the bias of uncorrected Hill estimates in the left tail. The blue upward pointing triangles (▲) show the bias of the corrected Hill estimates, using the cross-sectional mean, in the right tail. The red downward pointing triangles (▼) show the bias of the corrected Hill estimates, using the cross-sectional mean, in the left tail. The bars surrounding the mean estimates present the standard deviation of the estimates.