Technical appendix: Simulation results for GARCH processes and the approach using harmonic numbers

This technical appendix provides simulation results for GARCH processes and estimators of the tail index using harmonic numbers discussed in Section 3. Tables 5 and 6 present the numerical results on the performance of OLS estimators in regressions (1.3) with $\gamma = 0$ and $\gamma = 1/2$ for GARCH(1, 1) processes $Z_t = \sigma_t \epsilon_t$, where $\sigma_t^2 = \beta + \lambda Z_{t-1}^2 + \delta \sigma_{t-1}^2$, and ϵ_t are i.i.d. standard normal errors. The choice of the parameter values for β, λ and δ follows that in the simulation results presented by Kokoszka and Wolf (2004) who focused on subsampling approaches to estimating the mean of heavy-tailed observations. The corresponding values of the tail index ζ_0 of GARCH processes considered are provided in the same paper. The GARCH processes were simulated using the UCSD GARCH toolbox for Matlab by Kevin Sheppard. The IGARCH processes were simulated using the code by Mico Loretan.

Tables 7 and 8 provide simulation results for the Pareto exponent estimators in regression (3.10) for AR(1) and MA(1) processes driven by heavy-tailed innovations exhibiting deviations from power laws in form (4.14) and Student t distributions.

References

Kokoszka, P. and Wolf, M. (2004), 'Subsampling the mean of heavy-tailed dependent observations', *Journal of Time Series Analysis* **25**, 217–234.

Table 5. Behavior of the usual OLS estimator $\hat{b}_n$ in the regression
$\log(\text{Rank}) = a - b\log(\text{Size})$ for GARCH(1, 1) innovations

n	50	100	200	500
		Mean $\hat{b}_n$ (OLS s.e) (SD $\hat{b}_n$)		
$\beta = 1,\ \lambda = 1.3,$ $\delta = 0.05,\ \zeta_0 \approx 1.19$	1.375* (0.035) (0.417)	1.294* (0.019) (0.328)	1.234* (0.011) (0.252)	1.139 (0.006) (0.159)
$\beta = 1,\ \lambda = 1.1,$ $\delta = 0.1,\ \zeta_0 \approx 1.43$	1.587* (0.040) (0.465)	1.511* (0.022) (0.366)	1.411 (0.013) (0.276)	1.297* (0.007) (0.164)
$\beta = 1,\ \lambda = 0.9,$ $\delta = 0.15,\ \zeta_0 \approx 1.83$	1.926* (0.047) (0.534)	1.839* (0.026) (0.411)	1.761 (0.015) (0.305)	1.534 (0.010) (0.162)
$\beta = 1,\ \lambda = 0.9,$ $\delta = 0.1,\ \zeta_0 = 2$	2.057 (0.050) (0.530)	1.979 (0.028) (0.407)	1.881* (0.016) (0.296)	1.628* (0.011) (0.151)
$\beta = 1,\ \lambda = 0.5,$ $\delta = 0.5,\ \zeta_0 = 2$	2.315* (0.059) (0.665)	2.136* (0.033) (0.528)	1.983 (0.019) (0.391)	1.630* (0.013) (0.202)
$\beta = 1,\ \lambda = 0.1,$ $\delta = 0.9,\ \zeta_0 = 2$	3.799* (0.104) (0.880)	3.235* (0.060) (0.701)	2.677* (0.038) (0.546)	1.855* (0.022) (0.303)

Notes: The entries are the estimates of the tail index and their standard errors using regression (1.3) with $\gamma = 0$ for GARCH(1, 1) processes $Z_t = \sigma_t \epsilon_t$, where $\sigma_t^2 = \beta + \lambda Z_{t-1}^2 + \delta \sigma_{t-1}^2$, and ϵ_t are i.i.d. standard normal errors. "Mean $\hat{b}_n$" is the sample mean of the estimates $\hat{b}_n$ obtained in simulations, and "SD $\hat{b}_n$" is their sample standard deviation. "OLS s.e." is the OLS standard error in regression (1.3) with $\gamma = 0$. The value ζ_0 is the true tail index of Z_t. The asteric indicates rejection of the true null hypothesis $H_0 : \zeta = \zeta_0$ in favor of the alternative hypothesis $H_a : \zeta \neq \zeta_0$ at the 5% significance level using the reported OLS standard errors. The total number of observations $N = 2000$. Based on 10000 replications.

Table 6. Behavior of the usual OLS estimator $\hat{b}_n$ in the regression
$\log(\text{Rank} - 1/2) = a - b \log(\text{Size})$ for GARCH(1, 1) innovations

n	50	100	200	500
		Mean $\hat{b}_n^{\gamma=1/2}$ ($\sqrt{2/n}\times$Mean $\hat{b}_n^{\gamma=1/2}$) (SD $\hat{b}_n^{\gamma=1/2}$)		
$\beta = 1$, $\lambda = 1.3$, $\delta = 0.05$, $\zeta_0 \approx 1.19$	1.495 (0.299) (0.453)	1.366 (0.193) (0.346)	1.277 (0.128) (0.260)	1.159 (0.073) (0.162)
$\beta = 1$, $\lambda = 1.1$, $\delta = 0.1$, $\zeta_0 \approx 1.43$	1.727 (0.345) (0.505)	1.596 (0.226) (0.386)	1.492 (0.149) (0.285)	1.321 (0.084) (0.166)
$\beta = 1$, $\lambda = 0.9$, $\delta = 0.15$, $\zeta_0 \approx 1.83$	2.097 (0.419) (0.580)	1.943 (0.275) (0.432)	1.823 (0.182) (0.314)	1.562 (0.099) (0.164)
$\beta = 1$, $\lambda = 0.9$, $\delta = 0.1$, $\zeta_0 = 2$	2.243 (0.449) (0.585)	2.091 (0.296) (0.424)	1.951 (0.195) (0.308)	1.658* (0.105) (0.150)
$\beta = 1$, $\lambda = 0.5$, $\delta = 0.5$, $\zeta_0 = 2$	2.512 (0.502) (0.721)	2.271 (0.321) (0.555)	2.051 (0.205) (0.398)	1.658* (0.105) (0.203)
$\beta = 1$, $\lambda = 0.1$, $\delta = 0.9$, $\zeta_0 = 2$	4.116* (0.823) (0.948)	3.405* (0.482) (0.740)	2.745* (0.274) (0.566)	1.884 (0.119) (0.307)

Notes: The entries are the estimates of the tail index and their standard errors using regression (1.3) with $\gamma = 1/2$ for GARCH(1, 1) processes $Z_t = \sigma_t \epsilon_t$, where $\sigma_t^2 = \beta + \lambda Z_{t-1}^2 + \delta \sigma_{t-1}^2$, and ϵ_t are i.i.d. standard normal errors. "Mean $\hat{b}_n^{\gamma=1/2}$" is the sample mean of the estimates $\hat{b}_n^{\gamma}$ with $\gamma = 1/2$ obtained in simulations, and "SD $\hat{b}_n^{\gamma=1/2}$" is their sample standard deviation. The values $\sqrt{2/n}\times$Mean $\hat{b}_n^{\gamma=1/2}$ are the standard errors of $\hat{b}_n^{\gamma}$ with $\gamma = 1/2$ provided by Theorem 1. The value ζ_0 is the true tail index of Z_t. The asteric indicates rejection of the true null hypothesis $H_0 : \zeta = \zeta_0$ in favor of the alternative hypothesis $H_a : \zeta \neq \zeta_0$ at the 5% significance level using the reported standard errors. The total number of observations $N = 2000$. Based on 10000 replications.

Table 7. Behavior of the OLS estimator $\hat{b}'_n$ in the regression
$\log(H(t-1)) = a' - b' \log(\text{Size}_t)$ for innovations deviating from power laws

	n	50	100	200	500
AR(1)		Mean $\hat{b}'_n$			
c	ρ	($\sqrt{2/n}\times$Mean $\hat{b}'_n$) (SD $\hat{b}'_n$)			
0	0	1.002	0.998	0.995	0.996
		(0.200) (0.195)	(0.141) (0.140)	(0.100) (0.100)	(0.063) (0.062)
0	0.5	1.167	1.122	1.105	1.123
		(0.233) (0.318)	(0.159) (0.253)	(0.110) (0.201)	(0.071) (0.147)
0	0.8	1.462	1.337	1.266*	1.252*
		(0.292) (0.555)	(0.189) (0.435)	(0.127) (0.346)	(0.079) (0.269)
0.5	0	0.997	0.966	0.995	0.995
		(0.199) (0.194)	(0.141) (0.139)	(0.100) (0.099)	(0.063) (0.064)
0.5	0.5	1.161	1.120	1.105	1.122
		(0.232) (0.324)	(0.158) (0.249)	(0.110) (0.200)	(0.071) (0.149)
0.5	0.8	1.471	1.336	1.268*	1.257*
		(0.294) (0.557)	(0.189) (0.444)	(0.127) (0.345)	(0.080) (0.268)
0.8	0	1.004	0.995	0.996	0.995
		(0.201) (0.198)	(0.141) (0.138)	(0.100) (0.099)	(0.063) (0.063)
0.8	0.5	1.162	1.121	1.106	1.121
		(0.232) (0.324)	(0.159) (0.252)	(0.111) (0.199)	(0.071) (0.147)
0.8	0.8	1.475	1.340	1.266*	1.253*
		(0.295) (0.556)	(0.189) (0.436)	(0.127) (0.351)	(0.079) (0.268)
MA(1)		Mean $\hat{b}'_n$			
c	θ	($\sqrt{2/n}\times$Mean $\hat{b}'_n$) (SD $\hat{b}'_n$)			
0	0.5	1.066	1.047	1.039	1.052
		(0.213) (0.279)	(0.148) (0.201)	(0.104) (0.145)	(0.067) (0.095)
0	0.8	1.067	1.043	1.041	1.052
		(0.213) (0.294)	(0.147) (0.206)	(0.104) (0.149)	(0.067) (0.097)
0.5	0	0.999	0.996	0.995	0.995
		(0.200) (0.194)	(0.141) (0.140)	(0.100) (0.100)	(0.063) (0.063)
0.5	0.5	1.068	1.042	1.039	1.049
		(0.214) (0.277)	(0.147) (0.200)	(0.104) (0.143)	(0.066) (0.096)
0.5	0.8	1.075	1.049	1.043	1.051
		(0.215) (0.296)	(0.148) (0.211)	(0.104) (0.150)	(0.066) (0.098)
0.8	0	1.001	0.996	0.995	0.995
		(0.200) (0.196)	(0.141) (0.138)	(0.100) (0.099)	(0.063) (0.063)
0.8	0.5	1.068	1.045	1.042	1.049
		(0.214) (0.279)	(0.148) (0.197)	(0.104) (0.144)	(0.066) (0.095)
0.8	0.8	1.071	1.046	1.040	1.051
		(0.214) (0.291)	(0.148) (0.209)	(0.104) (0.148)	(0.066) (0.098)

Notes: The entries are estimates of the tail index and their standard errors using regression (3.10) for the AR(1) and MA(1) processes $Z_t = \rho Z_{t-1} + u_t$, $t \geq 1$, $Z_0 = 0$, and $Z_t = u_t + \theta u_{t-1}$, where i.i.d. u_t follow the distribution $P(Z > s) = s^{-\zeta}\left(1 + c(s^{-\alpha\zeta} - 1)\right)$, $s \geq 1$, with $\zeta = \alpha = 1$ and $c \in [0, 1)$. For a general case $\zeta > 0$, one multiplies all the numbers in the table by ζ. "Mean $\hat{b}'_n$" is the sample mean of the estimates $\hat{b}'_n$ obtained in simulations, and "SD $\hat{b}'_n$" is their sample standard deviation. The values $\sqrt{2/n}\times$Mean $\hat{b}'_n$ are the standard errors of $\hat{b}'_n$ provided by expansion (3.12). The asteric indicates rejection of the true null hypothesis $H_0 : \zeta = 1$ in favor of the alternative hypothesis $H_a : \zeta \neq 1$ at the 5% significance level using the reported standard errors. The total number of observations $N = 2000$. Based on 10000 replications.

Table 8. Behavior of the OLS estimator $\hat{b}_n$ in the regression $\log(H(t-1)) = a' - b' \log(\text{Size}_t)$ for Student t innovations

	n	50	100	200	500
AR(1)		Mean $\hat{b}'_n$			
m	ρ	($\sqrt{2/n}\times$Mean $\hat{b}'_n$) (SD $\hat{b}'_n$)			
2	0	1.959 (0.392) (0.370)	1.911 (0.270) (0.252)	1.827 (0.183) (0.165)	1.550* (0.098) (0.074)
2	0.5	2.153 (0.431) (0.488)	2.082 (0.295) (0.362)	1.995 (0.200) (0.253)	1.675* (0.106) (0.115)
2	0.8	2.634 (0.527) (0.843)	2.437 (0.345) (0.636)	2.253 (0.225) (0.443)	1.822 (0.115) (0.202)
3	0	2.763 (0.553) (0.501)	2.631 (0.372) (0.323)	2.417* (0.242) (0.194)	1.869* (0.118) (0.080)
3	0.5	3.077 (0.615) (0.629)	2.922 (0.413) (0.433)	2.683 (0.268) (0.270)	2.022* (0.128) (0.109)
3	0.8	3.921 (0.784) (1.103)	3.569 (0.505) (0.757)	3.141 (0.314) (0.463)	2.214* (0.140) (0.188)
4	0	3.409 (0.682) (0.588)	3.160 (0.447) (0.365)	2.820* (0.282) (0.204)	2.048* (0.130) (0.083)
4	0.5	3.813 (0.763) (0.706)	3.530 (0.499) (0.463)	3.116* (0.312) (0.266)	2.196* (0.139) (0.111)
4	0.8	4.897 (0.979) (1.168)	4.317 (0.610) (0.748)	3.617 (0.362) (0.428)	2.369* (0.150) (0.189)
MA(1)		Mean $\hat{b}_n$			
m	θ	($\sqrt{2/n}\times$Mean $\hat{b}'_n$) (SD $\hat{b}'_n$)			
2	0.5	2.097 (0.419) (0.480)	2.025 (0.286) (0.336)	1.935 (0.193) (0.224)	1.631* (0.103) (0.099)
2	0.8	2.141 (0.428) (0.565)	2.064 (0.292) (0.379)	1.962 (0.196) (0.249)	1.645* (0.104) (0.107)
3	0.5	3.002 (0.600) (0.620)	2.850 (0.403) (0.441)	2.605 (0.261) (0.253)	1.976* (0.125) (0.098)
3	0.8	3.156 (0.631) (0.752)	2.956 (0.418) (0.491)	2.677 (0.268) (0.290)	2.006* (0.127) (0.107)
4	0.5	3.715 (0.743) (0.691)	3.431 (0.485) (0.442)	3.038* (0.304) (0.255)	2.156* (0.136) (0.100)
4	0.8	3.943 (0.789) (0.847)	3.590 (0.508) (0.527)	3.128* (0.313) (0.296)	2.191* (0.139) (0.108)

Notes: The entries are estimates of the tail index and their standard errors using regression (3.10) for the AR(1) and MA(1) processes $Z_t = \rho Z_{t-1} + u_t$, $t \geq 1$, $Z_0 = 0$, and $Z_t = u_t + \theta u_{t-1}$, where i.i.d. u_t have the Student t distribution with m degrees of freedom. "Mean $\hat{b}'_n$" is the sample mean of the estimates $\hat{b}'_n$ obtained in simulations, and "SD $\hat{b}'_n$" is their sample standard deviation. The values $\sqrt{2/n}\times$Mean $\hat{b}'_n$ are the standard errors of $\hat{b}'_n$ provided by expansion (3.12). The asteric indicates rejection of the true null hypothesis $H_0 : \zeta = m$ in favor of the alternative hypothesis $H_a : \zeta \neq m$ at the 5% significance level using the reported standard errors. The total number of observations $N = 2000$. Based on 10000 replications.