

# A NEW REGRESSION-BASED TAIL INDEX ESTIMATOR

João Nicolau and Paulo M. M. Rodrigues\*

*Abstract*—A new regression-based approach for the estimation of the tail index of heavy-tailed distributions with several important properties is introduced. First, it provides a bias reduction when compared to available regression-based methods; second, it is resilient to the choice of the tail length used for the estimation of the tail index; third, when the effect of the slowly varying function at infinity of the Pareto distribution vanishes slowly, it continues to perform satisfactorily; and fourth, it performs well under dependence of unknown form. An approach to compute the asymptotic variance under time dependence and conditional heteroskedasticity is also provided.

## I. Introduction

OVER the past four decades, there has been considerable interest in the estimation of the tail index (henceforth,  $\alpha$ ) of heavy-tailed distributions in fields such as computer science, telecommunications, insurance, finance, and economics. The tail index is an indicator of the highest finite moment of a random variable and the rate at which the tail of a distribution decays. The smaller the value of  $\alpha$ , the lower the order of the highest finite moment and the slower the decay of the tail of the probability distribution will be, which increases the likelihood of extreme observations and outliers. Recent studies indicate that for stock returns in developed economies, typically  $\alpha \in (2, 5)$  (see Jansen & de Vries, 1991; Gabaix et al., 2006). For daily exchange rates of developed (Loretan & Phillips, 1994) and emerging markets (Ibragimov, Ibragimov, & Kattuman, 2013)  $\alpha \in (2, 4)$  has been observed, and in the context of income and wealth studies,  $\alpha \in (1.5, 3)$  has been reported (Gabaix, 2009; Ibragimov, Ibragimov, & Khamidov, 2011; for further examples, see Ibragimov, Ibragimov, & Walden, 2015). These empirical findings suggest that many variables typically analyzed may have a reduced number of finite moments.

The widespread interest in heavy-tailed distributions has led to the refinement and development of a number of tail index estimators over the years (see Hill, 1975; Kratz &

Resnick, 1996; Beirlant, Vynckier, & Teugels, 1996; Gabaix & Ibragimov, 2012. For reviews of these methods see Embrechts, Klüppelberg, & Mikosch, 2012, and Beirlant et al., 2004). Many of these procedures rely on plotting the statistic of interest against a number of the sample upper-order statistics and then inferring an appropriate value for  $\alpha$  from the properties of the resulting graph (Kratz & Resnick, 1996; Beirlant et al., 1996). Although many of these estimation methods exhibit interesting asymptotic properties, their finite sample performance is questionable (Kearns & Pagan, 1997; Huisman et al., 2001). Thus, the correct identification and estimation of  $\alpha$  remains a challenging and empirically relevant quest.

In this paper, we introduce a new regression-based procedure that overcomes several empirical difficulties of currently available tail index estimators. The new estimation method does not involve order statistics and can be applied in more general contexts than Pareto. There are three important features of our method, which can be summarized as follows: first, it provides a bias reduction when compared to regression-based methods (Gabaix & Ibragimov, 2012); second, it is more resilient to the choice of the number of largest observations used to estimate  $\alpha$  than the widely used Hill estimator; and third, when the effect of the slowly varying function at infinity of the Pareto distribution vanishes slowly, our estimator continues to perform satisfactorily, whereas the Hill estimator deteriorates rapidly.

The remainder of the paper is organized as follows. Section II briefly presents three frequently applied tail index estimators, which we will use for comparison with the new procedure proposed in this paper: section III introduces the new tail index estimator and discusses its asymptotic properties: section IV provides a discussion of a detailed Monte Carlo study of the finite sample bias of the tail index estimators discussed in sections II and III, as well as an analysis of the impact on these tests of empirically relevant features frequently found in economic and financial time series, such as time dependence and conditional heteroskedasticity. Section V illustrates the potential usefulness of the approach in an empirical application on absolute returns of 74 daily exchange rate series considering the U.S. dollar as numeraire; finally, section VI summarizes the main results. An online appendix provides detailed proofs of the results presented throughout the paper (appendix A), additional in-depth Monte Carlo analysis results (appendix B), and further estimation results of the empirical section (appendix C).

## II. Tail Index Estimators

In this section we briefly describe three procedures that are widely used in empirical work and will be used for comparison with the new approach introduced in section III. In what

Received for publication November 18, 2015. Revision accepted for publication June 5, 2018. Editor: Bryan S. Graham.

\*Nicolau: ISEG-Universidade de Lisboa and REM/CEMAPRE; Rodrigues: Banco de Portugal and NovaSBE, Universidade Nova de Lisboa.

We are grateful to two anonymous referees and the editor, Bryan S. Graham, for their helpful and constructive comments, as well as to Uwe Hasler for very useful discussions and insightful suggestions. We also thank the Exchange seminar participants at Banco de Portugal, seminar participants at Erasmus University Rotterdam, University of Barcelona, participants of the Workshop on Time Series at Goethe University Frankfurt, conference participants of the 9th annual meeting of the *Portuguese Economic Journal* and Karim Abadir, Giuseppe Cavaliere, Robin Lumsdaine, Morten Nielsen, Wing Wah Tham, Carlos Velasco, and Chen Zhou for useful comments and suggestions. This work was funded by Fundação para a Ciência e Tecnologia through project number UID/Multi/00491/2019, PTDC/EGE-ECO/28924/2017, as well as through UID/ECO/00124/2013, UID/ECO/00124/2019, and Social Sciences DataLab (LISBOA-01-0145-FEDER-022209), POR Lisboa (LISBOA-01-0145-FEDER-007722, LISBOA-01-0145-FEDER-022209), and POR Norte (LISBOA-01-0145-FEDER-022209).

A supplemental appendix is available online at [http://www.mitpressjournals.org/doi/suppl/10.1162/rest\\_a\\_00768](http://www.mitpressjournals.org/doi/suppl/10.1162/rest_a_00768).

follows, a heavy-tailed distribution is defined as a distribution function  $F$ , such that the survival function,  $\bar{F} := 1 - F$ , is regularly varying at infinity with index  $-\alpha$ ,

$$\bar{F}(x) := 1 - F(x) = x^{-\alpha} \mathcal{L}(x), \quad (1)$$

where  $0 < x < \infty$ ,  $\alpha > 0$  is a fixed unknown parameter and  $\mathcal{L}$  is a slowly varying function satisfying  $\lim_{t \rightarrow \infty} \frac{\mathcal{L}(tx)}{\mathcal{L}(t)} = 1$  for all  $x > 0$ .

A large number of tail index estimation procedures available in the literature are based on the largest order statistics  $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$  obtained from an independent and identically distributed (i.i.d.) sample  $\{X_t\}_{t=1}^n$  of a distribution function  $F$ . One such approach is the conditional maximum likelihood estimator (MLE) proposed by Hill (1975),

$$\hat{\alpha}_{Hill} := \left[ \frac{1}{m} \sum_{j=1}^m \log X_{(j)} - \log X_{(m+1)} \right]^{-1}, \quad (2)$$

where  $m := m(n) \rightarrow \infty$  is the number of largest-order statistics used in the estimation of  $\alpha$ .

The nice theoretical properties of this estimator, namely, consistency (Mason, 1982), asymptotic normality (Hall, 1982), and asymptotic efficiency (Beirlant, Bouquiaux, & Werker, 2006), have led researchers to develop improved variants and to show that these modifications work well in the Pareto case (see Drees, de Haan, & Resnick, 2000; Resnick & Stărică, 1997).

Available evidence suggests that the Hill estimator is most effective when the underlying distribution is either Pareto or close to Pareto (Drees et al., 2000). However, if the distribution is as in equation (1), the Hill estimator is only approximately MLE and its accuracy becomes less clear. Hence, although the Hill estimator is an interesting and powerful approach, it is not easily implemented empirically because of nuisance parameters whose feasible optimal choice is unknown but need to be specified for the adequate performance of the procedure.

**Remark 1.** The choice of  $m$  used in equation (2) has been an important topic of research, and several approaches for its determination in the i.i.d. context have been put forward (see Danielsson et al., 2001, and Nguyen & Samorodnitsky, 2012). This is an important concern since convergence in distribution of  $\hat{\alpha}_{Hill}$  critically depends on the rate at which the nuisance parameter  $m$  grows with the total sample size (see section IV).

**Remark 2.** A further shortcoming of this approach is the use of order statistics as these require sorting the data, which may become computationally expensive—at least  $O(n \log n)$  steps are required—and destroys the tie ordering of the data and their temporal structure (Stoev & Michailidis, 2010). Moreover, as Gabaix and Ibragimov (2012) noted, order statistics are dependent variables even for initial i.i.d. observations drawn from heavy-tailed power law distributions. Considering this feature, Gabaix and Ibragimov (2012) show that the

standard error of the OLS estimator of  $\alpha$  computed from a log-log rank-size regression (which will be discussed next) is asymptotically  $(2/n)^{1/2}\alpha$ . This is an important result, as the naive computation of OLS standard errors would significantly underestimate the true standard error of the OLS tail index estimator.

As a consequence of these shortcomings, simpler OLS regression-based estimation methods have attracted considerable attention among empirical researchers. One such alternative is the OLS-based log-log rank-size regression (see Rosen & Resnick, 1980, and Gabaix, 1999):

$$\log(t - \gamma) = \theta - \alpha \log X_{(t)} + \text{error}_t, \quad t = 1, \dots, m, \quad (3)$$

with  $\gamma = 0$ . This approach is based on the assumption that a nonnegative variable  $X$  has a power law distribution  $P(X \geq x) = Cx^{-\alpha}$ , for constant  $C > 0$  and  $\alpha > 0$ , which can be approximated by the linear relationship  $\log \frac{x}{n} \approx \log(C) - \alpha \log(X_{(t)})$  (see Gabaix & Ibragimov, 2012). The statistical properties of the OLS estimators of equation (3) have been analyzed in Gabaix and Ioannides (2004) and Gabaix and Ibragimov (2012). For reference purposes, we define the OLS estimator of  $\alpha$  computed from equation (3) as  $\hat{\alpha}_{\gamma=0}$ .

Finally, the third procedure we consider is by Gabaix and Ibragimov (2012), who propose estimation of equation (3) with the optimal shift of  $\gamma = 1/2$ . We denote the resulting estimator of  $\alpha$  as  $\hat{\alpha}_{\gamma=1/2}$ . The motivation for the proposal of this improved regression is that  $\hat{\alpha}_{\gamma=0}$  suffers from small sample bias. Considering i.i.d. random variables drawn from a Pareto distribution, Gabaix and Ibragimov (2012) show that using  $\gamma = 1/2$  instead of  $\gamma = 0$  in equation (3) significantly reduces this bias, while maintaining the good asymptotic properties of the estimator (see also the Monte Carlo results in section IV). Furthermore, according to Gabaix and Ibragimov's (2012) simulation results, the log-log rank-size regression estimator with  $\gamma = 1/2$  also performs well under dependence in the data, including GARCH-type processes exhibiting volatility clustering similar to real-world financial returns and under deviations from the exact power law distribution.

### III. The New Tail Index Estimator

#### A. The Estimator

To introduce the new tail index estimator, consider a sequence of i.i.d. random variables  $\{X_t\}_{t=1}^n$  drawn from a Pareto distribution with survival function,

$$\bar{F}(x) := P(X_t > x) = \left(\frac{x_0}{x}\right)^\alpha, \quad \text{with } x > x_0 > 0. \quad (4)$$

Applying logarithms to equation (4) and rearranging gives

$$\log \bar{F}_n(x_i) = \alpha \log x_0 - \alpha \log x_i + [\log \bar{F}_n(x_i) - \log \bar{F}(x_i)], \quad i = 1, 2, \dots \quad (5)$$

where  $\bar{F}_n(x_i) := \frac{1}{n} \sum_{t=1}^n I_{\{X_t > x_i\}}$  and  $I_{\{\cdot\}}$  is an indicator function.

Expression (5) can be seen as a regression equation where  $\log(\bar{F}_n(x_i))$  is the dependent variable,  $(\alpha \log x_0)$  is the intercept,  $(-\log x_i)$  is the explanatory variable, and  $[\log \bar{F}_n(x_i) - \log \bar{F}(x_i)] =: \varepsilon_i$  is the error term. Standard methods to estimate  $\alpha$  in a regression framework such as equation (3) treat  $x_i$  as an order statistics, say  $X_{(i)}$ , and as a result,  $\bar{F}_n(X_{(i)}) = i/n$ . Obviously, the order statistic  $X_{(i)}$  is a random variable, and the statistical properties of the estimation approach need to accommodate this feature. In contrast, the approach we propose treats  $x_i$  as a nonrandom variable, which considerably simplifies the analysis of the statistical properties of the resulting estimator.

To implement our procedure, we generate  $x_i$ , which is used to compute  $\log \bar{F}_n(x_i)$ , according to a deterministic scheme. We follow an importance sampling type approach, where instead of spreading the sample points of  $x_i$  out evenly, we concentrate the distribution of  $x_i$  in parts of the state-space of  $X$  that are of most ‘‘importance.’’ Since  $x_i := F^{-1}(u_i)$ , where  $F^{-1}$  is the quantile function of  $F$  and  $u_i \in (0, 1)$ , assuming that  $F$  is a Pareto distribution as in equation (4), we have that  $x_i = (1 - u_i)^{-\frac{1}{\alpha}} x_0$ . Note that we avoid treating  $u_i$  as a uniformly distributed random variable to keep  $x_i$  as a deterministic realization. Thus, we consider  $u_i := \frac{i}{m}$ ,  $i = 1, 2, \dots, m - 1$ .<sup>1</sup> Given that  $-\log x_i = \alpha^{-1} \log(1 - u_i) - \log x_0$ , we rule out  $F^{-1}(1)$  by imposing the upper limit of the index  $i$  to be at most  $m - 1$ . Consequently, equation (5) can be written as

$$y_i = \vartheta + \alpha z_i + \varepsilon_i, \quad i = 1, 2, \dots, m - 1, \tag{6}$$

where  $y_i := \log \bar{F}_n(x_i)$ ,  $x_i := (1 - u_i)^{-\frac{1}{\alpha}} x_0$ , and  $z_i := \alpha^{-1} \log(1 - u_i)$ .

Equation (6) is infeasible given the dependence of  $z_i$  on  $\alpha$ . To make equation (6) feasible, we deal with  $\alpha$  as a nuisance parameter and treat  $z_i$  as a generated regressor. The procedure consists of two steps. First, an initial estimate of  $\alpha$ ,  $\tilde{\alpha}$ , is considered to generate  $\tilde{z}_i = \tilde{\alpha}^{-1} \log(1 - u_i)$  (see remark 2). Second, the OLS estimate of  $\alpha$ , which we denote as  $\hat{\alpha}_{Pareto}$ , is computed from the feasible regression,

$$y_i = \vartheta + \alpha \tilde{z}_i + \varepsilon_i, \quad i = 1, 2, \dots, m - 1. \tag{7}$$

**Remark 3.** Replacing  $\alpha$  by  $\tilde{\alpha}$  in the first step has little impact on the estimation of  $\alpha$  from equation (7) since its effect vanishes asymptotically (see the proofs of theorems 1 and 2 in online appendix A). Moreover, it also follows from the proofs of theorems 1 and 2 that it is not necessary to assume  $\tilde{\alpha} \xrightarrow{P} \alpha$ . In fact, any value  $\tilde{\alpha} > 0$  leads to a consistent  $\hat{\alpha}_{Pareto}$  estimator, although choosing  $\tilde{\alpha}$  as an arbitrary value may originate efficiency losses. Thus, our estimate may be obtained

as follows. Set  $\tilde{\alpha} = c > 0$ , where  $c$  is any positive value, and generate  $\tilde{z}_i = \tilde{\alpha}^{-1} \log(1 - u_i)$ . Run regression (7) to obtain a new estimator  $\tilde{\alpha}^*$  and a new sequence of  $\tilde{z}_i^*$ , and run it again to obtain the final estimate  $\hat{\alpha}_{Pareto}$ . Our simulations have confirmed that this procedure ensures rapid convergence. This remark is important for two reasons. First, it indicates that the proposed estimator is self-sufficient in the sense that it can work alone, without the need to resort to any preliminary estimator; second, it simplifies the proof of the results, as we do not need to specify additional conditions on  $\tilde{\alpha}$ .

When  $X$  is governed by a strict Pareto distribution,  $x_0$  is the left end point of the support of  $X$  and can be estimated as  $\hat{x}_0 = \min(X_1, \dots, X_n)$ , the bias rapidly converges to 0, and the variance of  $\hat{\alpha}_{Pareto}$  is of order  $1/n$  (just as under MLE). One further interesting aspect of our methodology is that  $z_i$ , and to some extent  $\hat{z}_i$ , can be treated as a fixed explanatory variable. Moreover, the whole probabilistic structure of the model can be derived. In other words, considering the  $(m - 1) \times 1$  vector of errors  $\boldsymbol{\varepsilon} := (\varepsilon_1, \dots, \varepsilon_{m-1})'$ , it follows from Donsker’s CLT that  $\sqrt{n}\boldsymbol{\varepsilon} \xrightarrow{d} N(0, \boldsymbol{\Sigma}_{m-1})$ , where the covariance matrix  $\boldsymbol{\Sigma}_{m-1}$  is known exactly (no estimation is required; see lemma 1 and online appendix A for details). This is an important result as it allows us to obtain  $\text{Var}(\hat{\alpha}_{Pareto})$  and easily conduct generalized least squares estimation if required.

To extend the procedure to more general settings, we next consider Pareto-type tail behavior of the form

$$\bar{F}(x) = ax^{-\alpha} (1 + bx^{-\beta} + o(x^{-\beta})), \tag{8}$$

where  $a > 0$ ,  $\beta > \alpha$ , and  $b \in \mathbb{R}$ , which includes, among others, the nontrivial alpha-stable distribution and the Student- $t$  distribution. The parameters  $b$  and  $\beta$  govern the second-order behavior of the Taylor expansion and aim to reflect deviations from strict Pareto tail behavior. Although the specification in equation (8) is less general than equation (1), it is necessary for the theoretical analysis (see Hall, 1982).

Thus, following the same approach as in equation (6), we consider,

$$y_i = \vartheta + \alpha z_i + \tilde{\varepsilon}_i, \quad i = 1, 2, \dots, m - 1, \tag{9}$$

where  $y_i := \log \bar{F}_n(x_i)$ ,  $x_i := (1 - u_i)^{-\frac{1}{\alpha}} x_0$ , and  $z_i := \alpha^{-1} \log(1 - u_i)$ . The most important difference between equations (7) and (9) is that  $\tilde{\varepsilon}_i$  in the latter case is  $\tilde{\varepsilon}_i := \varepsilon_i + \eta_i$ , where  $\varepsilon_i := \log \bar{F}_n(x_i) - \log \bar{F}(x_i)$ ,  $\eta_i := \log(1 + b(1 - u_i)^{\frac{\beta}{\alpha}} x_0^{-\beta} + o(x_i^{-\beta}))$  and  $\bar{F}(x_i)$  is as defined in equation (8). The additional term,  $\eta_i$ , is responsible for the finite sample bias of our estimator, but this effect vanishes asymptotically under appropriate conditions (see the proof of theorem 3 in online appendix A). We show in section IV, that this finite sample bias is smaller than that of  $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{\gamma=0}$ , and  $\hat{\alpha}_{\gamma=1/2}$ .

As previously indicated, a crucial step of our method is the estimation of  $x_0$  (which is necessary to compute  $x_i = (1 - u_i)^{-\frac{1}{\alpha}} x_0$ ), which determines the window of values

<sup>1</sup>In the next section, we detail exactly how  $m$  relates to  $n$ , since it is of fundamental importance to understand the limit results of the new test statistics.



used to generate  $y_i = \log \bar{F}_n(x_i)$ ,  $i = 1, 2, \dots, m - 1$ . In this more general setting, we estimate  $x_0$  as  $\hat{x}_0 = \hat{F}^{-1}(1 - \kappa_n)$ ,  $0 < \kappa_n < 1$  ( $\hat{x}_0$  is the empirical quantile of order  $1 - \kappa_n$ ). The smaller the value of  $\kappa_n$ , the better the approximation of equation (8) to the tail of the Pareto law will be and, consequently, the closer  $\hat{\varepsilon}_i$  is to  $\varepsilon_i$ , that is, as  $\kappa_n \rightarrow 0$ , we have that  $\hat{x}_0 \rightarrow \infty$  and  $\hat{\varepsilon}_i \rightarrow \varepsilon_i$ .

*B. Asymptotic Properties*

*Preliminary results.* To characterize the asymptotic properties of our new estimator, consider first the results provided next in lemmas 1 to 5. In what follows, we consider either  $m$  or  $m_n$  in the computation of the relevant moments. This distinction is important as the first is used when no rate restriction is imposed on  $m$  with respect to  $n$  and  $m_n$  is considered otherwise. (Detailed proofs of all results in this section are available in online appendix A). Throughout this section, we use  $\xrightarrow{d}$  and  $\xrightarrow{p}$  to denote convergence in distribution and in probability, respectively, of a sequence of random elements when the sample size is allowed to diverge, and  $\lfloor \cdot \rfloor$  to represent the largest integer of the argument.

**Lemma 1.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of i.i.d. random variables with a continuous distribution function  $F(x)$ , and let  $\varepsilon := (\varepsilon_1, \dots, \varepsilon_{m-1})'$ , where  $\varepsilon_i := \log \bar{F}_n(x_i) - \log \bar{F}(x_i)$ ,  $m$  is some arbitrary, positive fixed integer, and  $x_0$  is known. Then, as  $n \rightarrow \infty$ , it follows, based on Donsker's CLT, that  $\sqrt{n}\varepsilon \xrightarrow{d} N(0, \Sigma_{m-1})$ , where  $N(0, \Sigma_{m-1})$  is a zero mean multivariate normal distribution with covariance matrix  $\Sigma_{m-1} := [\sigma_{ij}]_{(m-1) \times (m-1)}$ ,  $\sigma_{ij} := F(x_q)/\bar{F}(x_q)$ , and  $q := i \wedge j$ ,  $i, j = 1, 2, \dots, m - 1$ . If  $F_P$  is the Pareto distribution, with  $\bar{F}_P(x) := P(X_i > x) = (\frac{x_0}{x})^\alpha$  and  $x_i = (1 - u_i)^{-\frac{1}{\alpha}} x_0$ , then  $\sigma_{ij} := q/(m - q)$ .*

**Lemma 2.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of Pareto distributed i.i.d. random variables and consider  $\mathbf{w}_i := (1 \ z_i)'$ , with  $z_i := \alpha^{-1} \log(1 - i/m_n)$  and  $m_n = n$ . Then as  $n \rightarrow \infty$ , it follows from the extended continuous mapping theorem (van der Vaart & Wellner, 1996, theorem 1.11.1) that  $\frac{1}{m_n} \sum_{i=1}^{m_n-1} \mathbf{w}_i \sqrt{n}\varepsilon_i \xrightarrow{d} N(\mathbf{0}, \Omega)$ , where*

$$\text{Var} \left( \frac{1}{m_n} \sum_{i=1}^{m_n-1} \mathbf{w}_i \sqrt{n}\varepsilon_i \right) \rightarrow \Omega \text{ and } \Omega := \begin{bmatrix} 1 & -\frac{2}{\alpha} \\ -\frac{2}{\alpha} & \frac{5}{\alpha^2} \end{bmatrix}.$$

**Lemma 3.** *Let  $\mathbf{w}_i := [1 \ z_i]'$ , with  $z_i := \alpha^{-1} \log(1 - i/m)$ , where  $m$  is some arbitrary, positive integer. Then, as  $m \rightarrow \infty$ ,*

$$\frac{1}{m} \sum_{i=1}^{m-1} \mathbf{w}_i \mathbf{w}_i' \rightarrow \begin{bmatrix} 1 & -\frac{1}{\alpha} \\ -\frac{1}{\alpha} & \frac{2}{\alpha^2} \end{bmatrix}.$$

**Lemma 4.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of i.i.d. random variables with survival function given by equation (8). Considering  $m_n = \lfloor n^\theta \rfloor$  with  $\theta$  fixed and  $\frac{2}{2+\alpha} < \theta < 1$ , and  $\kappa_n := m_n/n = n^{\theta-1}$ , it follows that  $\hat{x}_0 = \hat{F}^{-1}(1 - \kappa_n) = \kappa_n^{-1/\alpha} x_0 + o_p(1)$ .*

**Lemma 5.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of i.i.d. random variables with survival function given by equation (8). Con-*

*sider  $\mathbf{w}_i := [1 \ z_i]'$  with  $z_i := \alpha^{-1} \log(1 - i/m_n)$ , and  $m_n = \lfloor n^\theta \rfloor$  with  $\theta$  fixed and  $\frac{2}{2+\alpha} < \theta < 1$ . Then, as  $n \rightarrow \infty$  (and consequently  $m_n \rightarrow \infty$ ),  $\frac{1}{\sqrt{m_n}} \sum_{i=1}^{m_n-1} \mathbf{w}_i \sqrt{n}\varepsilon_i \xrightarrow{d} N(0, \Omega^*)$ ,*

$$\text{Var} \left( \frac{1}{\sqrt{m_n}} \sum_{i=1}^{m_n-1} \mathbf{w}_i \sqrt{n}\varepsilon_i \right) \rightarrow \Omega^* \text{ and } \Omega^* := \begin{bmatrix} 2 & -\frac{3}{\alpha} \\ -\frac{3}{\alpha} & \frac{6}{\alpha^2} \end{bmatrix}.$$

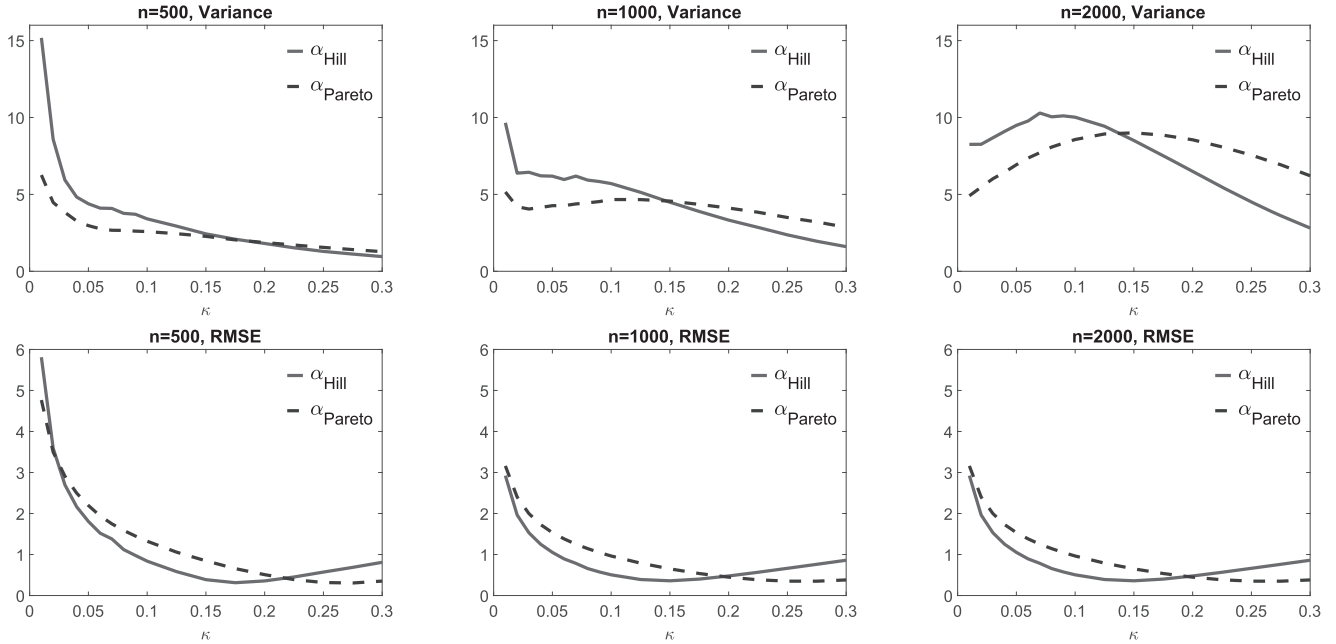
**Remark 4.** In lemma 2, we set  $m_n = n$  (although this could be relaxed), whereas in lemma 5,  $m_n$  has to verify the condition  $m_n = \lfloor n^\theta \rfloor$  with  $\frac{2}{2+\alpha} < \theta < 1$ . In both lemmas  $m_n \rightarrow \infty$ , but in the latter case,  $m_n$  grows more slowly than  $n$ . A complete justification can be provided only using asymptotic theory. Nevertheless, we note that in lemma 5, a Pareto-type distribution is considered, whereas in lemma 2, an exact Pareto is assumed. This is of fundamental importance since when a Pareto-type distribution is used, the grid of points  $\{\hat{x}_0, x_1, \dots, x_{m_n-1}\}$ , from which  $\bar{F}_n(x_i) := \frac{1}{n} \sum_{t=1}^n \mathbf{I}_{\{X_t > x_i\}}$  is computed, should start at  $\hat{x}_0 = \hat{F}^{-1}(1 - \kappa_n)$  where  $\kappa_n = m_n/n$ . In this case, a rate restriction on  $m_n$  with respect to  $n$  is required in order to guarantee two conditions. First,  $\kappa_n \rightarrow 0$  (hence,  $m_n/n \rightarrow 0$ ) needs to be ensured. In this way, the difference between the Pareto type and the exact Pareto distributions evaluated at  $x_i \in \{\hat{x}_0, x_1, \dots, x_{m_n-1}\}$  vanishes asymptotically. This is a consequence of the fact that the Pareto and the Pareto-type survival functions satisfy  $\bar{F}(xw)/\bar{F}(x) \rightarrow w^{-\alpha}$  as  $x \rightarrow \infty$  (see Beirlant et al., 2004, chap. 2). This condition is needed to guarantee consistency of the tail index estimator. Second, as shown in lemma 4,  $\hat{F}^{-1}(1 - \kappa_n)$  is a consistent estimator of the corresponding quantile function of the Pareto-type distribution, under the condition  $m_n = \lfloor n^\theta \rfloor$  with  $\frac{2}{2+\alpha} < \theta < 1$ . However, all of these conditions can be relaxed if the data come from an exact Pareto distribution if  $m_n \rightarrow \infty$ . In this case, the initial point  $x_0$  of the sequence  $\{\hat{x}_0, x_1, \dots, x_{m_n-1}\}$  can be determined as  $\hat{x}_0 = \min\{X_1, \dots, X_n\}$ , and  $\hat{F}^{-1}(1 - \kappa_n)$  and  $\kappa_n$  do not play any role.

*Properties of the tail index estimator under i.i.d.* The results in lemmas 1 to 5 allow us to state the following two theorems with the properties of the new estimator when data are generated by either a Pareto (theorem 1) or a Pareto-type distribution (theorem 2).

**Theorem 1.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of Pareto distributed i.i.d. random variables and set  $m_n = n$ . It follows from lemmas 1 to 3, as  $n \rightarrow \infty$ , that the OLS estimator computed from equation (7) is consistent,  $\hat{\alpha}_{\text{Pareto}} \xrightarrow{p} \alpha$ . In addition, if  $\tilde{\alpha} \xrightarrow{p} \alpha$ , then  $\sqrt{n}(\hat{\alpha}_{\text{Pareto}} - \alpha) \xrightarrow{d} N(0, 2\alpha^2)$ .*

**Theorem 2.** *Let  $\{X_i\}_{i=1}^n$  be a sequence of i.i.d. random variables with survival function as in equation (8). Assume that  $m_n = \lfloor n^\theta \rfloor$  with  $\theta$  fixed and  $\frac{2}{2+\alpha} < \theta < 1$ . Then, from lemmas 1, 3, 4, and 5, as  $n \rightarrow \infty$  (and consequently  $m_n \rightarrow \infty$ ), the OLS estimator computed from equation (8) is consistent,  $\hat{\alpha}_{\text{Pareto}} \xrightarrow{p} \alpha$ . In addition, if  $\tilde{\alpha} \xrightarrow{p} \alpha$ , then  $\sqrt{m_n}(\hat{\alpha}_{\text{Pareto}} - \alpha) \xrightarrow{d} N(0, 2\alpha^2)$ .*

FIGURE 1.—BEHAVIOR OF  $\hat{\alpha}_{Hill}$  AND  $\hat{\alpha}_{Pareto}$  IN THE PRESENCE OF AN IGARCH PROCESS WITH GAUSSIAN INNOVATIONS



The marginal distribution has a tail index of  $\alpha = 2$ . (Top) Ratio between the true and the theoretical variance of  $\hat{\alpha}_{Hill}$  and  $\hat{\alpha}_{Pareto}$ . (Bottom) RMSE of  $\hat{\alpha}_{Hill}$  and  $\hat{\alpha}_{Pareto}$ .

**Remark 5.** Theorem 1 establishes that the variance of  $\hat{\alpha}_{Pareto}$  is of order  $O_p(n^{-1})$ , where  $n$  represents the complete sample size. This order of convergence depends on the knowledge that the underlying distribution is exact Pareto. In a more general (and empirically relevant) setting, when the tail behavior is Pareto type as in equation (8), the variance of  $\hat{\alpha}_{Pareto}$  is of order  $O_p(m_n^{-1})$ , as referenced in theorem 2. Detailed proofs provided in online appendix A remarks A.2 and A.3 discuss how the different orders of convergence are obtained.

**Remark 6.** Interestingly, the proposed estimator reaches the same asymptotic variance as  $\hat{\alpha}_{v=1/2}$  proposed by Gabaix and Ibragimov (2012), with an important methodological difference: our estimator was obtained assuming Pareto-type tail behavior as in equation (8), whereas in Gabaix and Ibragimov (2012), only strict power law decline was considered, that is, imposing  $b = 0$ ,  $a = 1$ , and  $\sigma(x^b) = 0$  in equation (8).

**Remark 7.** Another important feature of our estimator is that its asymptotic variance does not depend on the parameters associated with the second-order behavior of the Taylor expansion, unlike several generalizations of the Hill estimator (see, e.g., Beirlant et al., 2004).

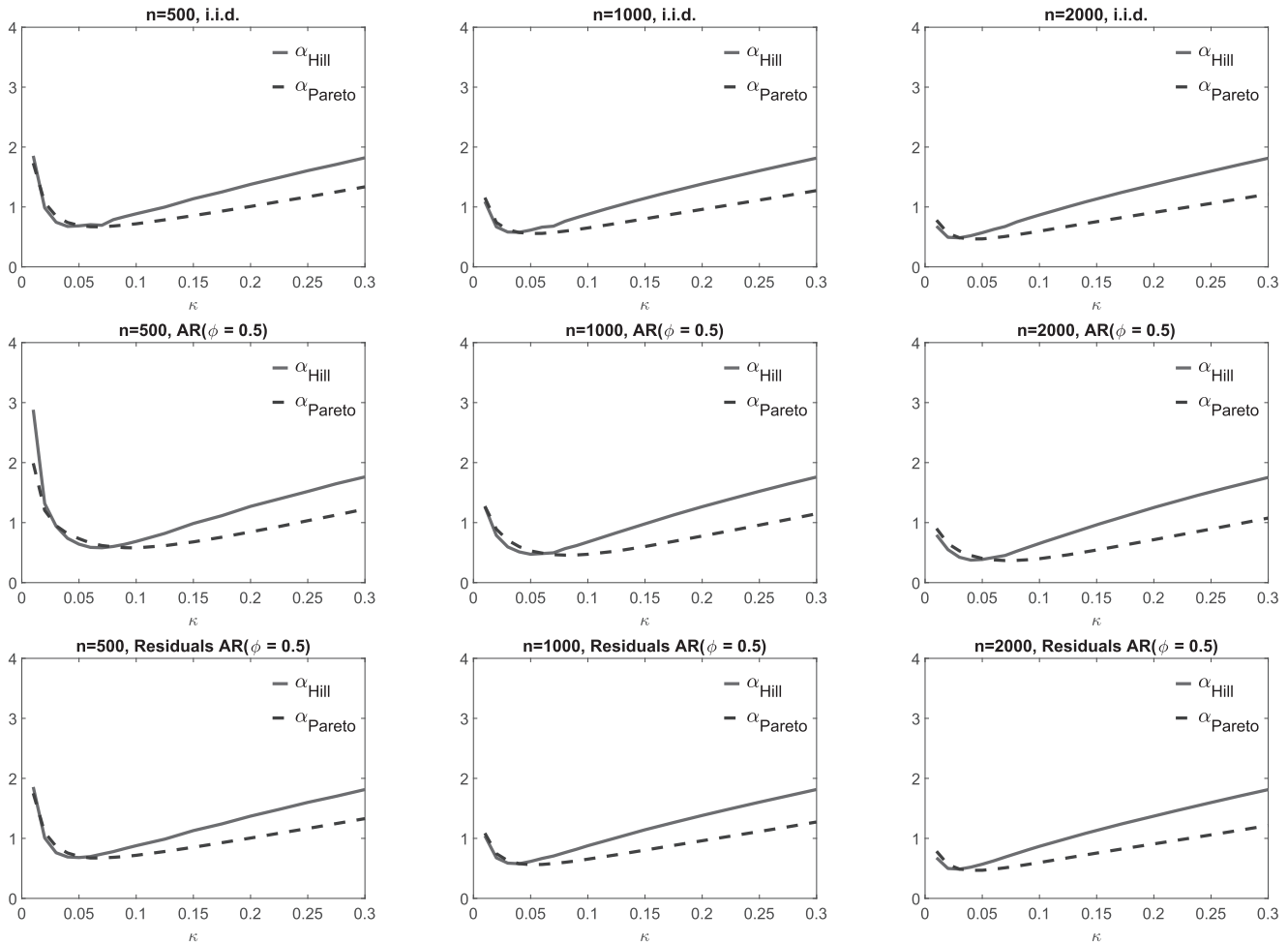
A practical issue, particularly in the context of Pareto-type distributions, relates to the determination of  $\kappa_n$ , which in turn determines  $m_n = \lfloor \kappa_n n \rfloor$  and  $x_0 = F^{-1}(1 - \kappa_n)$ . Based on our Monte Carlo analysis (see section IV and online appendix B), we recommend  $\kappa_n \leq 0.20$  for empirical applications. In general, the value of  $\kappa_n$  should shrink as the sample size increases and the level of dependence decreases (a detailed account of these relations is presented in section IV and in online appendix B). As our Monte Carlo study also reveals, an important feature of our procedure is that it is more robust

to the choice of  $\kappa_n$  than the Hill estimator. However, further research on the determination of the optimal  $\kappa_n$  is still needed.

*Properties of the tail index estimator under dependent data.* The results in theorems 1 and 2, derived in the i.i.d. context, are important since the proposed estimator is new to the literature and it was necessary that some of its basic properties were derived and compared with other well-known estimators. However, i.i.d. sequences have little relevance in most applications in economics and finance. For this reason, we now consider the case where  $\{X_t\}_{t=1}^n$  may exhibit dependence of unknown form.

We performed an extensive Monte Carlo analysis to address two specific questions (detailed results for all estimators discussed in this paper are available in online Appendix B): (a) How does dependence affect the asymptotic variance of  $\hat{\alpha}_{Pareto}$  (see figure 1), and (b) How does dependence affect the precision of  $\hat{\alpha}_{Pareto}$  and  $\hat{\alpha}_{Hill}$  (see figure 2). Our main conclusions are that, first, dependence in the data, via AR or GARCH dynamics, has important impacts on the limiting distributions of  $\hat{\alpha}_{Hill}$  and  $\hat{\alpha}_{Pareto}$ . However, there are significant differences between AR and GARCH dependence. Autocorrelation has a moderate effect on the asymptotic variance of  $\hat{\alpha}_{Pareto}$  and decreases as the sample size and  $m_n$  increase. This impact vanishes completely when the original process is replaced by the residuals of a suitable AR model. The limiting distribution of  $\hat{\alpha}_{Hill}$  is more affected, especially when  $m_n$  is small, but like  $\hat{\alpha}_{Pareto}$ , the autocorrelation effect tends to decline as the sample size and  $m_n$  increase. In contrast, GARCH effects have a strong impact on the asymptotic variance of both estimators (for results, on  $\hat{\alpha}_{Hill}$ , see Hill, 2010), especially when  $m_n$  is in the range of “optimal values.” Some of our conclusions are

FIGURE 2.—RMSE OF  $\hat{\alpha}_{Hill}$  AND  $\hat{\alpha}_{Pareto}$



DGP: Student- $t$  distribution with  $\alpha = 3$ .

aligned with the results of Kearns and Pagan (1997). Second, AR or GARCH dependence has an impact on the optimal choice of  $\kappa_n$  (and consequently  $m_n$ ): the larger the dependence, the larger is the optimal value of  $\kappa_n$ . The change in the optimal choice of  $\kappa_n$  is even more marked in the IGARCH case. Interestingly, dependence in the data does not seem to affect the quality of  $\hat{\alpha}_{Hill}$  and  $\hat{\alpha}_{Pareto}$  as long as  $\kappa_n$  is properly determined. Another important aspect is that  $\hat{\alpha}_{Pareto}$  performs better than  $\hat{\alpha}_{Hill}$  for most values of  $\kappa_n$  considered.

In sum, our findings point to a valid tail index estimator,  $\hat{\alpha}_{Pareto}$ , but to an inconsistent estimator of its variance when the i.i.d. hypothesis is wrongly assumed under dependence. This is an issue that is also common to the other estimators described in section II.<sup>2</sup> Hence, in what follows we develop asymptotic results for the variance of  $\hat{\alpha}_{Pareto}$  under general conditions.

<sup>2</sup>The problems with inconsistent standard errors under (nonlinear) dependence are somewhat similar to the problems with the use of sample autocorrelations and autocorrelation functions in the analysis of data with volatility clustering and GARCH-type processes (see, for instance, Mikosch & Starica, 2000). We thank a referee for pointing this out.

A careful analysis of the proofs of theorems 1 and 2 shows that the random quantity  $\frac{1}{m_n} \sum_{i=1}^{m_n-1} w_i \varepsilon_i$ , where  $\varepsilon_i := \log \bar{F}_n(x_i) - \log \bar{F}(x_i)$ , is the crucial element to discuss the consistency and the limiting distribution of our estimator. As such, some of the conditions imposed below have ultimately to do with convergence in probability and distribution of the empirical process  $\bar{F}_n(x)$  for dependent data.

Let  $\mathcal{F}_a^b$  denote the  $\sigma$ -fields generated by the random variables  $X_a, \dots, X_b$ , and define  $a(n) := \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty, k \geq 1\}$ . Thus, the process  $\{X_t\}_{t=1}^n$  is strongly mixing if  $a(n) \rightarrow 0$ .

**Theorem 3.** *Let  $\{X_t\}_{t=1}^n$  be a strictly stationary process with distribution  $F$  and survival function given by equation (8) and assume that  $\{X_t\}_{t=1}^n$  satisfies the strong mixing condition  $a(n) \rightarrow 0$ . Furthermore, assume that  $m_n = \lfloor n^\theta \rfloor$  with  $\theta$  fixed and  $\frac{2}{2+\alpha} < \theta < 1$ . (a) It follows, as  $n \rightarrow \infty$ , that the OLS estimator computed from equation (9) is consistent,  $\hat{\alpha}_{Pareto} \xrightarrow{P} \alpha$ ; (b) In addition, if  $a(n) = o(n^{-6-\delta})$ ,  $\delta \in (0, 1)$ , and  $\tilde{\alpha} \xrightarrow{P} \alpha$ , then  $\sqrt{m_n}(\hat{\alpha}_{Pareto} - \hat{\alpha}) \xrightarrow{d} N(0, V_{22})$ , where  $V_{22}$  is the*

equation (2) element of  $\mathbb{V} := \mathbb{E}\mathbb{J}\mathbb{E}$ , with  $\mathbb{E} := \begin{bmatrix} 2 & \alpha \\ \alpha & \alpha^2 \end{bmatrix}$ ,  
 $\mathbb{J} := \lim_{m_n \rightarrow \infty} \mathbf{J}_{m_n}$ ;  $\mathbf{J}_{m_n} := \frac{1}{m_n} \sum_{s=1}^{m_n-1} \sum_{i=1}^{m_n-1} \lim_{n \rightarrow \infty} E(\mathbf{V}_s \mathbf{V}_i')$ ,  
 $\mathbf{V}_i := \mathbf{w}_i \varepsilon_i$  and  $\mathbf{w}_i := (1 \ z_i)'$ .

**Remark 8.** Note that the conditions required in theorem 3 are broad enough to include many processes of interest in economics and finance (e.g., ARMA, GARCH, Markov-Switching processes, among others; see, for instance, Stelzer, 2009).

To estimate  $\mathbf{J}_{m_n}$  in the presence of autocorrelation and heteroskedasticity of unknown form, we consider a HAC estimator (as in Andrews & Monahan, 1992), which is based on the following procedure. Consider  $\hat{\boldsymbol{\theta}} := (\hat{\vartheta} \ \hat{\alpha})'$  a  $\sqrt{m_n}$ -consistent estimator of  $\boldsymbol{\theta}$ . First, we estimate a  $b$ th-order VAR model for  $\mathbf{V}_t = \mathbf{w}_t \hat{\varepsilon}_t$  ( $2 \times 1$  vector),  $\mathbf{V}_t = \sum_{i=1}^b \hat{\mathbf{A}}_i \mathbf{V}_{t-i} + \hat{\mathbf{v}}_t$ ,  $t = b + 1, \dots, m_n$ , where  $\hat{\mathbf{v}}_t$  is the corresponding residual vector. Second, we compute a standard kernel-based HAC estimator, say  $\hat{\mathbf{J}}_{m_n}^*(\hat{S}_{m_n})$ , based on the VAR residual vector  $\hat{\mathbf{v}}_t$ , that is,  $\hat{\mathbf{J}}_{m_n}^*(\hat{S}_{m_n}) = \frac{m_n}{m_n-2} \sum_{i=-m_n+1}^{m_n-2} \mathbf{K}\left(\frac{j}{\hat{S}_{m_n}}\right) \hat{\Gamma}_j^*$ , where  $\hat{\Gamma}_j^* := \begin{cases} \frac{1}{m_n} \sum_{t=j+1}^{m_n-1} \hat{\mathbf{v}}_t \hat{\mathbf{v}}_{t-j}' & \text{for } j \geq 0 \\ \frac{1}{m_n} \sum_{t=-j+1}^{m_n-1} \hat{\mathbf{v}}_{t+j} \hat{\mathbf{v}}_t' & \text{for } j < 0 \end{cases}$ ,  $\mathbf{K}(x) := \frac{25}{12\pi^2 x^2} \left( \frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right)$  is a real-valued kernel and  $\hat{S}_{m_n} := 1.3221 (\hat{c} m_n)^{1/5}$  is a data-dependent bandwidth parameter, with  $\hat{c} = \frac{4\hat{\rho}^2 \hat{\sigma}^4}{(1-\hat{\rho})^8} / \frac{\hat{\sigma}^4}{(1-\hat{\rho})^8}$ , and  $\rho$  and  $\sigma^2$  denoting the autoregressive and innovation variance parameters of the second component of  $\mathbf{V}_t$ , respectively, that is,  $\{z_t \hat{\varepsilon}_t\}$ . Third, one recolors the estimator  $\hat{\mathbf{J}}_{m_n}^*(\hat{S}_{m_n})$  to obtain the VAR prewhitened kernel estimator of  $\mathbf{J}_{m_n}$ , namely  $\hat{\mathbf{J}}_{m_n}(\hat{S}_{m_n}) = \hat{\mathbf{D}} \hat{\mathbf{J}}_{m_n}^*(\hat{S}_{m_n}) \hat{\mathbf{D}}'$ , with  $\hat{\mathbf{D}} := \left( \mathbf{I}_2 - \sum_{i=1}^b \hat{\mathbf{A}}_i \right)^{-1}$ .

**Theorem 4.** Considering that  $\sqrt{m_n}(\hat{\mathbf{A}}_i - \mathbf{A}_i) = O_p(1)$ ,  $i = 1, \dots, b$ , where  $b$  corresponds to the VAR order,  $\left( \mathbf{I}_2 - \sum_{i=1}^b \mathbf{A}_i \right)$  is nonsingular and the conditions of theorem 3 hold, then  $\hat{\mathbf{J}}_{m_n}(\hat{S}_{m_n}) - \mathbf{J}_{m_n} \xrightarrow{p} \mathbf{0}$ .

#### IV. Monte Carlo Analysis

A Monte Carlo study is conducted to assess and compare the root mean square errors (RMSE) of  $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{\gamma=0}$ ,  $\hat{\alpha}_{\gamma=1/2}$ , and  $\hat{\alpha}_{Pareto}$  as a function of  $\kappa_n$ . As DGPs, we consider several heavy-tailed distributions that satisfy equation (8), such as the Student- $t$ , Burr, and alpha-stable distributions,<sup>3</sup> which are

<sup>3</sup>Data from an alpha-stable distribution with index  $\alpha$  are generated as  $X_t = \frac{\sin(\alpha \mu_t)}{|\cos(\mu_t)|^{1/\alpha}} \left( \frac{\cos((1-\alpha)\mu_t)}{z_t} \right)^{\frac{1-\alpha}{\alpha}}$  where  $\mu_t$  is uniform on  $(-\pi/2, \pi/2)$  and  $z_t$  is an exponential variate with mean 1 (Samorodnitsky & Taqqu, 1994).

frequently used in the extreme-value literature (see Beirlant et al., 2004).<sup>4</sup>

The Monte Carlo experiments consist of the following steps. First, generate a sample of size  $n$ , with  $n \in \{500, 2000\}$ , from a heavy-tailed distribution. Second, fix a value for  $\kappa$  from the set  $\kappa \in \{0.01, 0.02, 0.03, \dots, 0.09, 0.1, 0.125, 0.15, \dots, 0.375, 0.4\}$ . The parameter  $\kappa$  is defined as an additive sequence of fixed increments of 0.01 from 0.01 to 0.1 and of 0.025 from 0.1 henceforward. Third, estimate  $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{\gamma=0}$ ,  $\hat{\alpha}_{\gamma=1/2}$ , and  $\hat{\alpha}_{Pareto}$  from a subsample of size  $m_n = \lfloor \kappa n \rfloor$ . Fourth, using the same distribution (and the same parameters), repeat steps 1 and 3 1,000 times and compute  $RMSE(\hat{\alpha}_i) := \sqrt{\sum_{t=1}^{1000} (\hat{\alpha}_{i,t} - \alpha)^2} / 1,000$ , where  $i$  refers to one of the four estimators under analysis. Fifth, select another value for  $\kappa$  in the above set and repeat all previous steps. Sixth, select another DGP and/or value for  $\alpha$  and repeat all previous steps.

Figure 3 compares the performance of  $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{\gamma=0}$ , and  $\hat{\alpha}_{\gamma=1/2}$ .

Three main conclusions are shown in figure 3. First, the minimum RMSE of  $\hat{\alpha}_{Hill}$  and  $\hat{\alpha}_{\gamma=1/2}$  are approximately equal, although they are generally attained at different values of  $\kappa$ . Second, in contrast to the regression-based estimators, the RMSE of the Hill estimator rapidly deteriorated as  $\kappa$  moves away from its minimum; as a consequence,  $\hat{\alpha}_{\gamma=0}$  and  $\hat{\alpha}_{\gamma=1/2}$  are more resilient to variations in  $\kappa$ . Third,  $\hat{\alpha}_{\gamma=1/2}$  displays better performance than  $\hat{\alpha}_{\gamma=0}$  for most values of  $\kappa$ , and there is significant bias reduction when  $\hat{\alpha}_{\gamma=1/2}$  is used comparatively to  $\hat{\alpha}_{\gamma=0}$ , as was also indicated by Gabaix and Ibragimov (2012). Thus, these results show that  $\hat{\alpha}_{\gamma=1/2}$  is in general the best among the three estimators we have just analyzed.

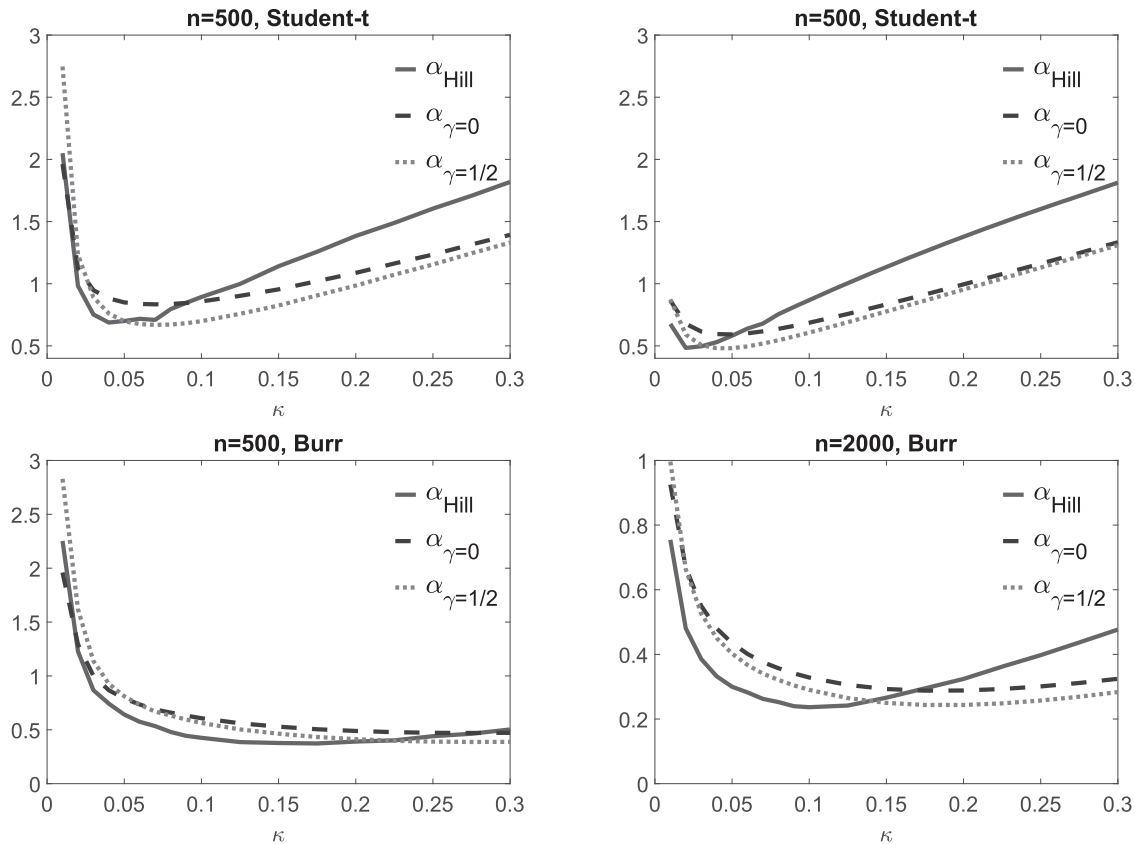
Given the dominant performance of  $\hat{\alpha}_{\gamma=1/2}$ , in figure 4 we compare this estimator to  $\hat{\alpha}_{Pareto}$  through the ratio  $\frac{RMSE(\hat{\alpha}_{\gamma=1/2})}{RMSE(\hat{\alpha}_{Pareto})}$  under various values for  $\kappa$ , sample sizes, and DGPs.

This figure shows that for most cases considered,  $\frac{RMSE(\hat{\alpha}_{\gamma=1/2})}{RMSE(\hat{\alpha}_{Pareto})} > 1$ , suggesting that the new estimator introduced presents an interesting performance when compared to  $\hat{\alpha}_{\gamma=1/2}$ . Specifically, we note that although in our simulations,  $\kappa$  is allowed to reach a value of 0.3, in general, and as recommended earlier in section IIIB, a  $\kappa \leq 0.2$  is typically used in the literature—for example,  $\kappa = 0.05$  or  $\kappa = 0.1$ . However, the Monte Carlo results also show (see online appendix B) that under persistent GARCH dependence, larger values for  $\kappa$  may be required.

Figure 4 indicates that the approach introduced obtains a smaller RMSE for the Burr distribution with  $T = 500$  and  $T = 2,000$  and for the alpha-stable distribution with  $T = 500$ . For the Student- $t$  and the alpha-stable distribution with  $T = 2000$ ,  $\hat{\alpha}_{Pareto}$  has similar performance to  $\hat{\alpha}_{\gamma=1/2}$  and for a Student- $t$  distribution with  $T = 500$ ,  $\hat{\alpha}_{\gamma=1/2}$  has superior performance to  $\hat{\alpha}_{Pareto}$ . More results are provided in online appendix B.

<sup>4</sup>Further results based on a large number of other DGPs are presented in online appendix B.



FIGURE 3.—PERFORMANCE OF  $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{\gamma=0}$  AND  $\hat{\alpha}_{\gamma=1/2}$ 

DGPs: Student- $t$  and Burr distributions with  $\alpha = 3$ .

## V. Empirical Application

Exchange rate dynamics has received considerable attention in the literature. In the context of studies on the heavy-tail properties of these series, particularly for developed markets, several important contributions have been made (see Hols & de Vries, 1991; Koedijk et al., 1992; Loretan & Phillips, 1994; Ibragimov, Dovidova, & Khamidov, 2010; Hartmann, Straetmans, & de Vries, 2010; Ibragimov et al., 2013). Emerging markets have received less attention (Ibragimov et al., 2013). We look to reduce this gap by observing the tail properties of daily absolute returns of 74 exchange rate series, considering the U.S. dollar (USD) as base currency, of which 59 belong to emerging markets. The currencies considered are classified into two groups (see table 1).

Detailed results on the point estimates of the tail index of the daily absolute returns, standard errors, and confidence intervals for all 74 series are provided in online appendix C. Figure 5 presents point estimates of the tail index for all 74 countries, with the 59 EM classified by region. The period of analysis is January 1, 1999, to May 16, 2016, and all data are obtained from Datastream. The tail indices for DM and EM are estimated using the four estimators discussed in the text ( $\hat{\alpha}_{Hill}$ ,  $\hat{\alpha}_{Pareto}$ ,  $\hat{\alpha}_{\gamma=0}$ , and  $\hat{\alpha}_{\gamma=1/2}$ ). The Hill estimator is computed based on 5% and 10% truncation levels ( $\kappa$ ), whereas for the other estimators, 10% and 15% truncation levels are

used. For the sake of space, figure 5 presents results only for  $\hat{\alpha}_{Hill}$  obtained with  $\kappa = 0.05$  and for  $\hat{\alpha}_{Pareto}$ ,  $\hat{\alpha}_{\gamma=0}$  and  $\hat{\alpha}_{\gamma=1/2}$  with  $\kappa = 0.10$ . The choice of the truncation levels follows from the Monte Carlo results of section IV, which indicate that the minimum RMSE for the estimators, in the case of volatility in the data (as is the case in the series considered), is obtained with a higher truncation level. In online appendix C, we report results for all the  $\kappa$  values used in the estimation of the tail indices.

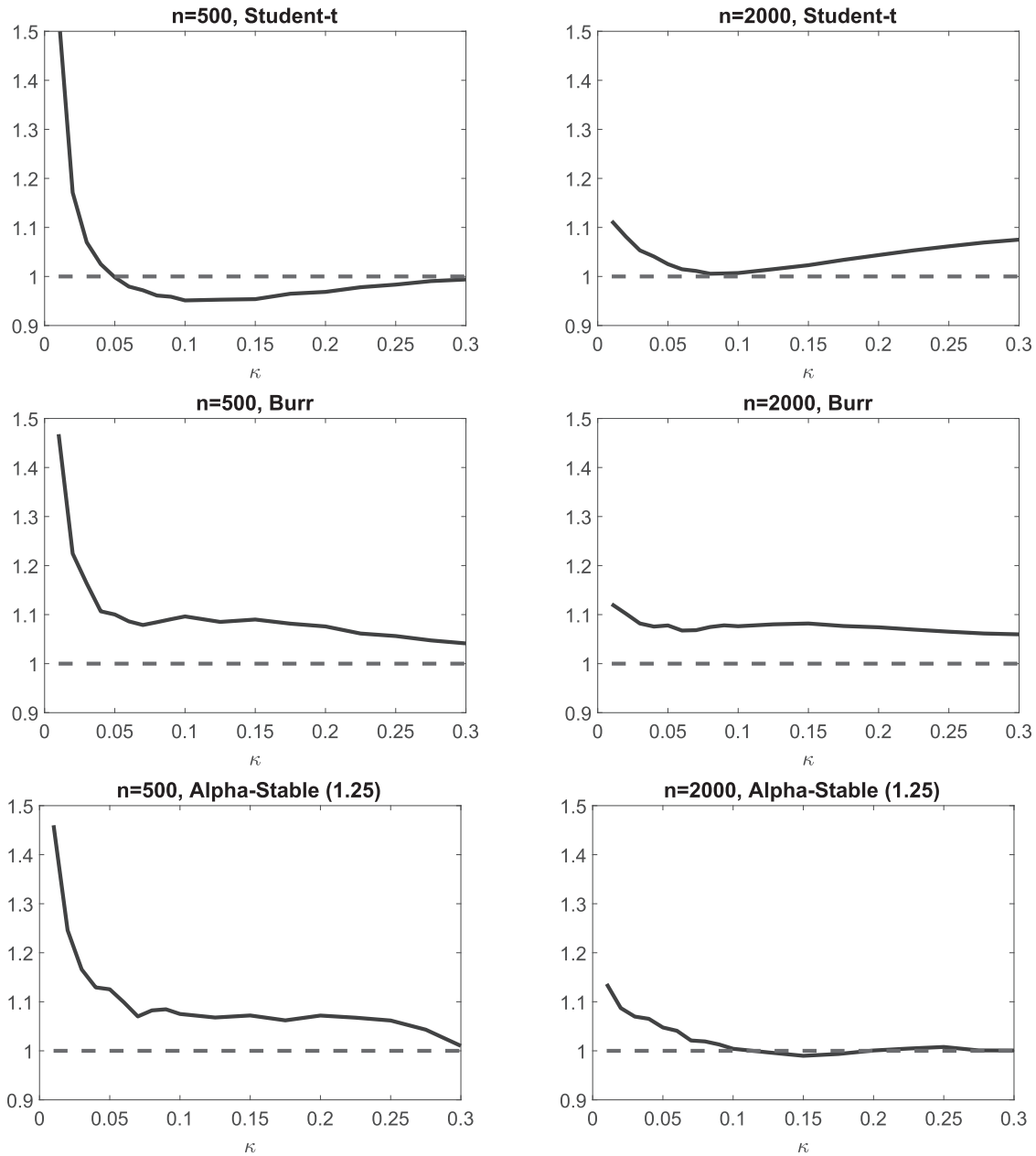
Figures 5 and C.2 (see also tables C.2 to C.4 in online appendix C) show that the tail indices of the exchange rate returns series computed in the complete sample (January 1, 1999–May 16, 2016) and in the two subperiods (subperiod I: January 1, 1999–September 15, 2008, and subperiod II: September 16, 2008–May 16, 2016), show considerable difference for the two groups of countries under analysis.

### A. Complete Sample

In the complete sample, 15 of the 59 EM display point estimates of the tail index below 2 ( $\hat{\alpha} < 2$ ) (see figure 5 and table 2). In particular, the results for  $\hat{\alpha}_{Pareto}$  indicate that the tail index for DM lies between [2.27, 4.27], whereas for EM, we observe that in Asia, the estimates are between [1.08, 3.28]—the lowest corresponding to the Saudi riyal



FIGURE 4.—RATIO  $RMSE(\hat{\alpha}_{\nu=1/2})/RMSE(\hat{\alpha}_{Pareto})$



DGPs: Student-*t* and Burr distributions with  $\alpha = 3$ , and alpha-stable distribution with  $\alpha = 1.25$ .

(SAR) and the highest to the Philippine peso (PHP). In Africa, it is between [1.27, 4.37], with the lowest corresponding to the Guinea frank (GNF) and the highest to the Tunisian dinar (TND). For Latin America, it is between [1.41, 3.42], with the lowest corresponding to the Argentina peso (ARS) and the largest to the Chilean peso (CLP). And for Europe, the range is [1.70, 4.26], where the lowest is obtained for the Ukrain hryvnia (UAH) and the highest for the Bulgarian lev (BGN).

Table 2 classifies the tail index estimates according to their statistical significance. The results reported and the corresponding 95% confidence intervals are computed based on

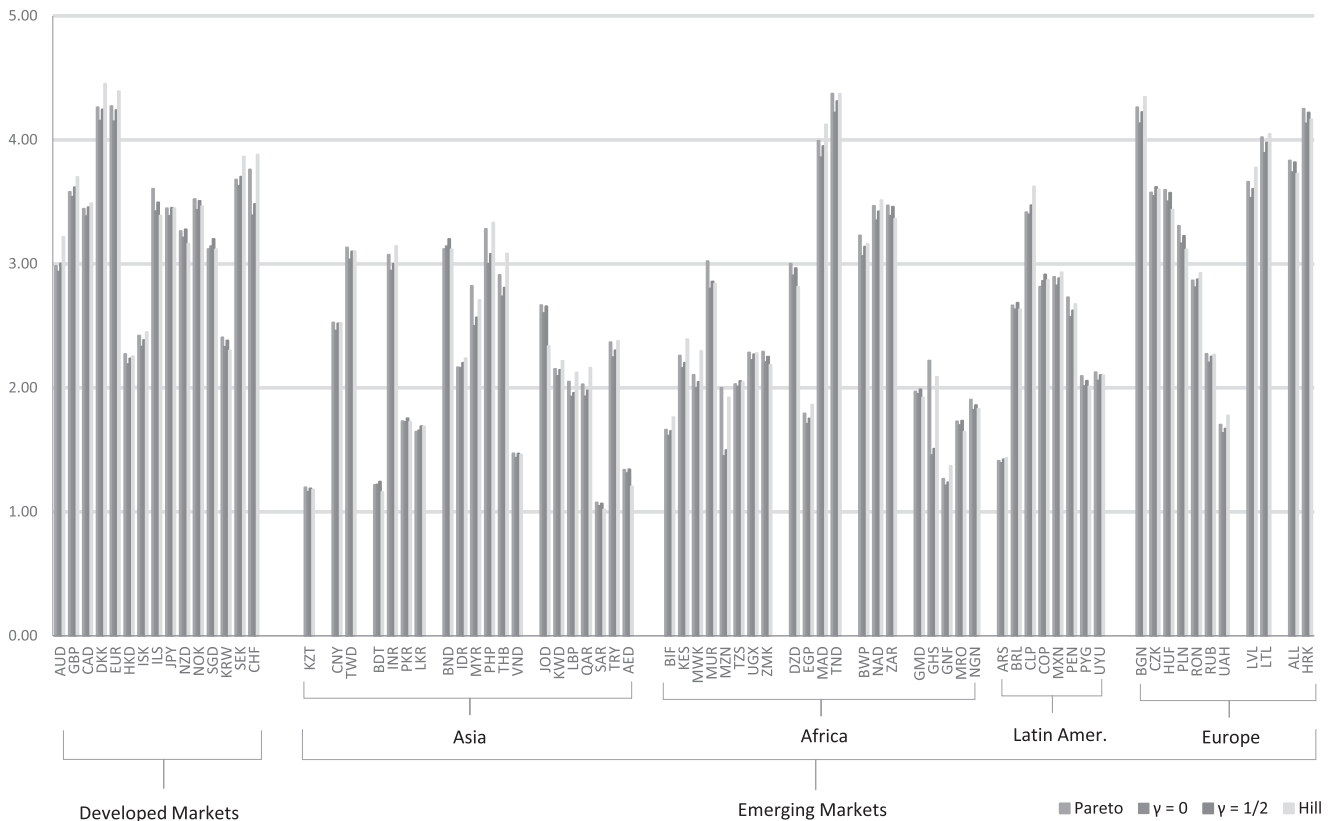
$\hat{\alpha}_{Pareto}$  and the robust standard error estimators discussed in section III. Overall, the tail index estimates of the 74 series lie between [1, 4]. Specifically, the tail indices belonging to DM lie between [2, 4] and those of EM between [1, 4]. Around 39% (29 exchange rates) of the series analyzed display statistically significant tail indices between [1, 2] suggesting that these series display infinite variances and one (SAR) appears not even to have finite first-order moments.<sup>5</sup> Forty-five of the

<sup>5</sup>Extreme heavy-tailedness, such as that observed for SAR, raises several problems: (a) the conclusions of many economic and financial models reverse, (b) diversification becomes suboptimal, (c) applications of

TABLE 1.—CURRENCY ABBREVIATIONS OF THE 74 DEVELOPED AND EMERGING MARKETS EXCHANGE RATE SERIES

Developed Markets (DM)	Emerging Markets (EM) Asia	Africa	Latin America	Europe
Australian dollar (AUD)	Kazakhstan tengge (KZT)	Burundi franc (BIF)	Argentine peso (ARS)	Bulgarian lev (BGN)
Canadian dollar (CAD)	Chinese yuan (CNY)	Kenyan shilling (KES)	Brazilian real (BRL)	Czech koruna (CZK)
Swiss franc (CHF)	Taiwan dollar (TWD)	Malawi kwacha (MWK)	Chilean peso (CLP)	Hungarian forint (HUF)
Danish krone (DKK)	Bangladeshi taka (BDT)	Mauritius rupee (MUR)	Colombian peso (COP)	Polish zloty (PLN)
Euro (EUR)	Indian rupee (INR)	Mozambique new metical (MZN)	Mexican peso (MXN)	Romanian new lei (RON)
Great Britain pound (GBP)	Pakistan rupee (PKR)	Tanzanian shilling (TZS)	Paraguay guarani (PYG)	Russian ruble (RUB)
Hong Kong dollar (HKD)	Sri Lanka rupee (LKR)	Uganda shilling (UGX)	Peruvian Nuevo sol (PEN)	Ukraine hryvnia (UAH)
Iceland krona (ISK)	Brunei dollar (BND);	Zambian kwacha (ZMK)	Uruguayan peso (UYU)	Latvian lats (LVL)
Israeli New shekel (ILS)	Indonesian rupiah (IDR)	Algerian dinar (DZD)		Lithuanian litas (LTL)
Japanese yen (JPY)	Malaysian ringgit (MYR)	Egyptian pound (EGP)		Albanian lek (ALL)
South Korean won (KRW)	Philippine peso (PHP)	Moroccan dirham (MAD)		Croatian kuna (HRK)
New Zealand dollar (NZD)	Thai baht (THB)	Tunisian dinar (TND)		
Norwegian kroner (NOK)	Vietnamese dong (VND)	Botswana pula (BWP)		
Singapore dollar (SGD)	Jordanian dinar (JOD)	Namibia dollar (NAD)		
Swedish krona (SEK)	Kuwaiti dinar (KWD)	South African rand (ZAR)		
	Lebanese pound (LBP)	Gambian dalasi (GMD)		
	Qatari rial (QAR)	Ghanaian new cedi (GHS)		
	Saudi riyal (SAR)	Guinea franc (GNF)		
	Turkish new lira (TRY)	Mauritanian ouguiya (MRO)		
	United Arabs Emirates dirham (AED)	Nigerian naira (NGN)		

FIGURE 5.—TAIL INDEX POINT ESTIMATES,  $\hat{\alpha}_k, k = \text{Pareto}, \gamma = 0, \gamma = 1/2, \text{Hill}$ , OF EXCHANGE RATE RETURNS, JANUARY 1, 1999–MAY 16, 2016.



series have finite second-order moments (series with  $\hat{\alpha} > 2$ ), 19 series have finite third-order moments ( $\hat{\alpha} > 3$ ), and none of the series seems to display finite fourth-order moments. The results, though having considered a larger period of analysis

(1999–2016), are generally in conformity with the results in tables 5 and 6 of Ibragimov et al. (2013).<sup>6</sup>

To further illustrate the adequacy of the tail truncation levels used, in particular for  $\hat{\alpha}_{\text{Pareto}}$ , figure C.1 in appendix C

standard risk measures such as value at risk and expected shortfall may become problematic, and (d) can originate difficulties in the application of standard econometric methods (see, e.g., Ibragimov et al. (2015).

<sup>6</sup>Note that Ibragimov et al. (2013) analyze only nine DM currencies (AUD, CAD, CHF, DKK, EUR, GBP, JPY, NOK, and SEK) and nine EM currencies (CNY, HKD, INR, KRW, MYR, RUB, SGD, THB, and TWD).

TABLE 2.—TAIL INDEX ESTIMATES,  $\hat{\alpha}_{Pareto}$ , OF THE ABSOLUTE EXCHANGE RATE RETURNS AND CORRESPONDING 95% CONFIDENCE INTERVALS FOR THE COMPLETE SAMPLE, AND SUBPERIODS I AND II

Full Sample												
$\alpha = 1$	$1 < \alpha < 2$	$\alpha = 2$	$2 < \alpha < 3$	$\alpha = 3$	$3 < \alpha < 4$	$\alpha = 4$	$4 < \alpha < 5$	$\alpha = 5$				
SAR = 1.08 [0.89, 1.26]	ARS = 1.41 [1.22, 1.60]	MRO = 1.73 [1.47, 1.99]	EGP = 1.79 [1.49, 2.10]	LBP = 2.05 [0.98, 3.12]	UGX = 2.28 [1.86, 2.70]	BRL = 2.66 [2.28, 3.05]	AUD = 2.98 [2.50, 3.47]	JPY = 3.45 [2.88, 4.02]	PHP = 3.28 [2.72, 3.85]	CAD = 3.44 [2.97, 3.92]	ALL = 3.83 [3.30, 4.37]	HUF = 3.60 [2.95, 4.25]
AED = 1.34 [1.12, 1.56]	PKR = 1.73 [1.39, 2.07]	GHS = 2.22 [1.94, 2.50]	MWK = 2.10 [1.79, 2.46]	CNY = 2.53 [2.04, 3.01]	UYU = 2.13 [1.79, 2.46]	ISK = 2.42 [2.07, 2.77]	BND = 3.12 [2.78, 3.67]	MUR = 3.02 [2.49, 3.30]	PLN = 3.31 [2.64, 3.98]	CLP = 3.42 [2.95, 3.89]	BGN = 4.26 [3.52, 5.01]	ILS = 3.61 [2.90, 4.31]
BDT = 1.22 [0.99, 1.44]	VND = 1.47 [1.23, 1.72]	GMD = 1.97 [1.64, 2.31]	MZN = 2.00 [1.67, 2.33]	ZMK = 2.29 [1.86, 2.73]	ISK = 2.42 [2.07, 2.77]	BWP = 3.23 [2.78, 3.67]	MXN = 2.87 [2.49, 3.30]	RON = 2.87 [2.46, 3.28]	CHF = 3.58 [3.05, 4.11]	GBP = 3.58 [3.04, 4.48]	LTL = 4.02 [3.32, 4.72]	
BIF = 1.66 [1.37, 1.96]	UAH = 1.70 [1.43, 1.98]	HKD = 2.27 [1.92, 2.63]	NGN = 1.90 [1.60, 2.21]	NGN = 1.90 [1.60, 2.21]	KRW = 2.41 [2.06, 2.76]	COP = 2.82 [2.37, 3.26]	MYR = 3.58 [2.36, 3.29]	SGD = 3.12 [2.57, 3.67]	LVL = 3.66 [3.00, 4.33]	CZK = 3.58 [3.01, 4.14]	MAD = 3.99 [3.29, 4.70]	
GNF = 1.27 [1.09, 1.45]	IDR = 2.17 [1.79, 2.54]	PYG = 2.10 [1.78, 2.41]	PYG = 2.10 [1.78, 2.41]	PYG = 2.10 [1.78, 2.41]	RUB = 2.27 [1.95, 2.60]	DZD = 3.00 [2.41, 3.59]	NAD = 3.47 [2.91, 4.03]	THB = 2.91 [2.50, 3.32]	NOK = 3.52 [3.04, 4.00]	DKK = 4.26 [3.59, 4.94]	SEK = 3.68 [3.03, 4.33]	
KZT = 1.20 [1.02, 1.37]	KES = 2.26 [1.82, 2.70]	QAR = 2.03 [1.59, 2.47]	QAR = 2.03 [1.59, 2.47]	QAR = 2.03 [1.59, 2.47]	TRY = 2.37 [2.10, 2.64]	INR = 3.07 [2.38, 3.77]	NZD = 3.26 [2.75, 3.78]	TWD = 3.13 [2.63, 3.63]	ZAR = 3.47 [2.98, 3.97]	EUR = 4.27 [3.59, 4.95]	TND = 4.37 [3.54, 5.21]	
LKR = 1.65 [1.40, 1.90]	KWD = 2.15 [1.83, 2.48]	TZS = 2.03 [1.75, 2.31]	TZS = 2.03 [1.75, 2.31]	TZS = 2.03 [1.75, 2.31]	JOD = 2.67 [1.90, 3.43]	PEN = 2.73 [2.17, 3.29]	PEN = 2.73 [2.17, 3.29]	PEN = 2.73 [2.17, 3.29]	HRK = 4.25 [3.55, 4.95]	HRK = 4.25 [3.55, 4.95]	HRK = 4.25 [3.55, 4.95]	
Subperiod I												
$\alpha = 1$	$1 < \alpha < 2$	$\alpha = 2$	$2 < \alpha < 3$	$\alpha = 3$	$3 < \alpha < 4$	$\alpha = 4$	$4 < \alpha < 5$	$\alpha = 5$				
BDT = 1.14 [0.88, 1.41]	AED = 1.31 [1.06, 1.56]	EGP = 2.12 [1.73, 2.51]	MYR = 1.94 [1.48, 2.40]	TRY = 2.19 [1.81, 2.57]	BRL = 2.43 [1.96, 2.90]	BWP = 3.17 [2.57, 3.77]	KRW = 3.25 [2.36, 4.14]	RON = 2.58 [2.06, 3.10]	ALL = 4.26 [3.30, 5.22]	LTL = 4.49 [3.63, 5.34]	CAD = 4.54 [4.01, 5.07]	BGN = 4.91 [4.37, 5.46]
GNF = 1.03 [0.84, 1.22]	ARS = 1.44 [1.15, 1.74]	GHS = 2.14 [1.72, 2.56]	MZN = 1.73 [1.40, 2.05]	TZS = 2.35 [1.88, 2.82]	IDR = 2.39 [1.86, 2.92]	CNY = 2.81 [2.19, 3.43]	KWD = 2.69 [1.96, 3.41]	TWD = 2.89 [2.33, 3.44]	AUD = 3.95 [3.07, 4.84]	LVL = 4.03 [3.28, 4.77]		DKK = 5.01 [4.46, 5.56]
SAR = 0.91 [0.69, 1.13]	BIF = 1.37 [1.10, 1.63]	HKD = 2.20 [1.75, 2.65]	NGN = 1.77 [1.43, 2.12]	UAH = 1.77 [1.43, 2.12]	MWK = 2.50 [1.92, 3.07]	COP = 2.65 [2.09, 3.20]	MXN = 3.54 [2.70, 4.38]		BND = 3.92 [3.22, 4.61]	NOK = 4.41 [3.50, 5.31]		EUR = 4.98 [3.83, 6.14]
GMD = 1.68 [1.36, 2.00]	JOD = 2.00 [1.53, 2.47]	PKR = 1.60 [1.11, 2.10]	UGX = 2.23 [1.81, 2.66]	UGX = 2.23 [1.81, 2.66]	THB = 2.58 [2.10, 3.06]	DZD = 2.96 [2.56, 3.37]	NAD = 3.35 [2.51, 4.45]		CHF = 4.58 [3.78, 5.39]	PLN = 4.04 [3.28, 4.80]		GBP = 5.12 [4.71, 5.53]
LBP = 1.69 [1.34, 2.04]	KES = 2.02 [1.55, 2.50]	PYG = 1.88 [1.48, 2.28]	PYG = 1.88 [1.48, 2.28]	PYG = 1.88 [1.48, 2.28]	ILS = 3.48 [2.51, 4.45]	NZD = 3.92 [2.74, 5.10]			CLP = 3.92 [3.16, 4.69]	SEK = 4.60 [3.89, 5.31]		HRK = 4.96 [3.98, 5.94]
LKR = 1.71 [1.38, 2.03]	KZT = 1.81 [1.40, 2.22]	QAR = 1.88 [1.12, 2.64]	QAR = 1.88 [1.12, 2.64]	QAR = 1.88 [1.12, 2.64]	INR = 2.76 [2.26, 3.27]	PEN = 2.67 [2.14, 3.19]			CZK = 4.54 [3.22, 4.61]	SGD = 3.92 [3.22, 4.61]		MAD = 4.74 [4.00, 5.48]
MRO = 1.37 [1.04, 1.70]	MUR = 2.20 [1.57, 2.84]	RUB = 1.73 [1.37, 2.09]	ZMK = 2.38 [1.82, 2.94]	ZMK = 2.38 [1.82, 2.94]	ISK = 3.05 [2.35, 3.75]	PHP = 2.75 [2.13, 3.38]			HUF = 4.47 [3.67, 5.27]	ZAR = 3.78 [3.32, 4.24]		TND = 4.98 [4.47, 5.49]
Subperiod II												
$\alpha = 1$	$1 < \alpha < 2$	$\alpha = 2$	$2 < \alpha < 3$	$\alpha = 3$	$3 < \alpha < 4$	$\alpha = 4$	$4 < \alpha < 5$	$\alpha = 5$				
KZT = 1.00 [0.74, 1.25]	AED = 1.44 [1.06, 1.81]	VND = 1.43 [1.13, 1.73]	BDT = 1.70 [1.27, 2.13]	ISK = 1.93 [1.46, 2.41]	NGN = 1.93 [1.43, 2.44]	IDR = 2.44 [1.92, 2.95]	AUD = 2.66 [2.09, 3.23]	COP = 2.87 [2.18, 3.55]	MAD = 3.89 [2.82, 4.96]	NZD = 2.90 [2.28, 3.52]	TRY = 2.85 [2.26, 3.43]	ALL = 3.86 [3.09, 4.62]
ARS = 1.44 [1.13, 1.75]	BIF = 2.22 [1.51, 2.92]	JOD = 2.70 [1.57, 4.84]	PKR = 2.09 [1.60, 2.58]	JOD = 2.70 [1.57, 4.84]	PKR = 2.09 [1.60, 2.58]	RUB = 2.54 [2.01, 3.08]	BND = 3.05 [2.26, 3.85]	CZK = 3.32 [2.54, 4.09]	MUR = 3.41 [2.28, 4.54]	PEN = 2.61 [2.00, 3.21]	TWD = 3.19 [2.21, 4.18]	BGN = 4.01 [3.72, 4.30]
EGP = 1.47 [1.24, 1.71]	CNY = 2.52 [1.88, 3.16]	KES = 2.52 [1.47, 3.57]	PYG = 2.39 [1.81, 2.96]	KES = 2.52 [1.47, 3.57]	PYG = 2.39 [1.81, 2.96]	PYG = 2.39 [1.81, 2.96]	BRL = 2.89 [2.26, 3.51]	DZD = 2.76 [2.05, 3.47]	MXN = 2.80 [2.14, 3.46]	PLN = 3.44 [2.89, 4.00]	UYU = 2.71 [2.01, 3.40]	DKK = 3.96 [3.01, 4.92]
LKR = 1.60 [1.20, 2.00]	GHS = 2.25 [1.86, 2.64]	KRW = 2.26 [1.66, 2.85]	QAR = 2.65 [1.07, 5.13]	KRW = 2.26 [1.66, 2.85]	QAR = 2.65 [1.07, 5.13]	QAR = 2.65 [1.07, 5.13]	BWP = 3.18 [2.80, 3.55]	GBP = 3.03 [2.41, 3.65]	MYR = 3.61 [2.43, 4.79]	RON = 3.23 [2.51, 3.95]	ZAR = 3.23 [2.59, 3.87]	EUR = 3.97 [2.99, 4.95]
MRO = 1.66 [1.31, 2.02]	GMD = 2.10 [1.66, 2.54]	KWD = 2.05 [1.59, 2.51]	TZS = 1.79 [1.39, 2.19]	GMD = 2.10 [1.66, 2.54]	KWD = 2.05 [1.59, 2.51]	TZS = 1.79 [1.39, 2.19]	CAD = 3.24 [2.50, 3.98]	ILS = 3.53 [2.68, 4.39]	MZN = 2.64 [1.95, 3.34]	SEK = 3.61 [2.79, 4.42]		HRK = 3.93 [3.66, 4.20]
SAR = 1.24 [0.89, 1.59]	GNF = 1.70 [1.31, 2.10]	LBP = 2.22 [1.39, 3.04]	UGX = 2.34 [1.81, 2.86]	LBP = 2.22 [1.39, 3.04]	UGX = 2.34 [1.81, 2.86]	UGX = 2.34 [1.81, 2.86]	CHF = 3.41 [2.16, 4.15]	INR = 3.14 [2.30, 4.53]	NAD = 3.23 [2.59, 3.87]	SGD = 3.05 [2.26, 3.85]		HUF = 3.70 [3.13, 4.27]
UAH = 1.72 [1.42, 2.03]	HKD = 2.38 [1.71, 3.04]	MWK = 1.87 [1.45, 2.29]	ZMK = 2.14 [1.66, 2.63]	HKD = 2.38 [1.71, 3.04]	MWK = 1.87 [1.45, 2.29]	ZMK = 2.14 [1.66, 2.63]	CLP = 3.03 [2.41, 3.65]	JPY = 3.02 [2.24, 3.80]	NOK = 3.46 [2.78, 4.14]	THB = 4.11 [2.41, 5.81]		LVL = 4.02 [3.17, 4.86]

The definitions of the currency acronyms are provided in section V and in the accompanying online appendix C. The shaded cells represent the currencies belonging to DM.

presents plots of the point estimates for EUR, CNY, GBP, and RUB, obtained with this estimator and the corresponding 95% confidence intervals for different values of the truncation level  $\kappa$ . This figure highlights the stability of the point estimates across the different truncation levels considered, as the confidence intervals for different values of  $\kappa$  intersect (this is also consistent with results in Ibragimov et al., 2015, pp. 96–97, who provide similar plots for  $\hat{\alpha}_{\gamma=1/2}$ ).<sup>7</sup>

Currency crises have generally been a characteristic of the international monetary system, and as such, detailed analysis of the tail index by subperiods may reveal further insights. Hence, we reestimate the tail indices over the two subperiods previously indicated: subperiod I (January 1, 1999–September 15, 2008) and subperiod II (September 16, 2008–May 16, 2016). The ranges of values of the tail indices of the 74 series in the two subperiods is [1, 5] for subperiod I and [1, 4] for subperiod II, which suggest different exchange rate dynamics in these two subperiods (see table 1).

### B. Subperiod I

Subperiod I includes the massive default of Argentina's external debt, the consequent abandoning of its currency board, and the devaluation of its currency in early 2002. Moreover, the initial part of subperiod I possibly still reflects the aftermath of the Asian financial crisis in 1997–1998 and the Russian financial crisis of 1998. However, from 2002 until the financial disruption in 2008, DM and EM overall did not experience a period of crises.

The point estimates in table 1 and figure C.2 in online appendix C indicate a large group of exchange rate return series, essentially belonging to EM, with an  $\hat{\alpha} \leq 2$ . The results show that the point estimates of the tail index for DM are between [2.20, 5.12], whereas for EM, the estimates for Asia are between [0.91, 3.92], the lowest corresponding to the Saudi riyal (SAR) and the highest to the Brunei dollar (BND). For Africa, these are in the range [1.03, 4.98], where the lowest corresponds to the Guinea frank (GNF) and the highest to the Tunisian dinnar (TND). For Latin America, these are between [1.44, 3.92], where the lowest corresponds to the Argentina peso (ARS) and the largest to the Chilean peso (CLP). Finally for Europe, the range of values is [1.73, 4.96], where the lowest is obtained for the Russian ruble (RUB) and the highest for the Croatian kuna (HRK).

In terms of statistical significance, we observe that for 31 of the 74 series,  $\hat{\alpha} \leq 2$  (which suggests that 42% of the series do not seem to have finite second moments); 41 series display  $\hat{\alpha} > 2$ ; 23 series present  $\hat{\alpha} > 3$ ; and only 8 series display  $\hat{\alpha} > 4$ . Hence, although in this period  $\hat{\alpha} \in [1, 5]$ , the distribution of the tail index estimates is very asymmetric as a large

concentration of estimates is observed between [1, 2] (see table 1).

### C. Subperiod II

Subperiod II includes the global financial crisis of 2008–2009, the European sovereign debt crisis, and a number of geopolitical events (e.g., the Arab Spring), which forced sharp depreciations in many DM as well as EM. Interestingly, the financial contagion originated by the global financial crisis was short, and by 2009, many EM had recovered access to the international financial system at low interest rates. This was likely the consequence of the change to flexible managed floating regimes made by EM (see Reinhart & Rogoff, 2004), as well as a result of accumulation of foreign exchange reserves.

The results for  $\hat{\alpha}_{Pareto}$  (see figure C.2 of online appendix C. indicate that the tail index for DM in this period is between [1.93, 3.97], whereas for EM, the estimates for Asia are between [1.00, 4.57], the lowest corresponding to the Kazakhstan tense (KZT) and the highest to the Philippine peso (PHP); in Africa, between [1.47, 4.11], where the lowest corresponds to the Egyptian pound (EGP) and the highest to the Tunisian dinnar (TND); in Latin America, the interval is [1.44, 3.03], with the lowest estimate corresponding to the Argentina peso (ARS) and the largest to the Chilean peso (CLP); and for Europe, we have [1.72, 4.02] where the Ukraine hryvnia (UAH) displays the lowest tail index and the Latvian lats (LVL) the highest.

In terms of statistical significance, 30 of the 74 series display  $\hat{\alpha} \leq 2$ , which corresponds to almost the same number of series with infinite variance as in subperiod I. However, twelve (BDT, BIF, GMD, GNF, LBP, MUR, MYR, MZN, RUB, SAR, TRY, and UYU) of the currencies that displayed an  $\hat{\alpha} \leq 2$  in subperiod I registered an increase in magnitude of their tail index in subperiod II, whereas four (KZT, EGP, VND, and UAH) see their tail index estimate decrease to  $\hat{\alpha} \leq 2$  in this period.

Specifically, we observe from table 1 that with the exception of HKD, ILS, and NZD, whose tail indices appear unchanged, all other DM currencies' tail index decreased. For EM, we note that 16 of the 59 currencies' tail index decreased, 28 remained unchanged, and 15 increased. These results suggest that there are differences between the tail index estimates in the two subperiods analyzed. In order to confirm this, we use the  $\hat{\alpha}_{Pareto}$  estimates and the corresponding robust standard errors computed in each subperiod and test the equality of the estimated tail indices in both subperiods. The results indicate statistically significant differences in 24 of the 74 series analyzed. In particular, 11 of the tail indices (BDT, BIF, GNF, MUR, MYR, MZN, PHP, RUB, THB, TRY, and UYU) are larger in subperiod II than in subperiod I, suggesting a potential reduction of the risk profile, whereas 13 (AUD, BGN, CAD, CHF, DKK, EGP, GBP, HRK, ISK, KRW, KWD, KZT, and MWK) decreased in subperiod II

<sup>7</sup>Plots for all exchanges rates and different sample periods analyzed are available from the authors.



compared to subperiod I (7 of which belonging to DM as expected).

#### D. Discussion

It is well documented in the literature that trades made by large market participants (such as central banks) can originate volatility and that their potentially large price impacts explain certain power law behavior (Gabaix et al., 2006; Gaibaix & Maggiori, 2015). Gabaix et al. (2003) show that the tail behavior of absolute returns of a stock price,  $|r_t|$ , is well characterized by  $P(|r_t| > x) \sim x^{-\alpha}$  with  $\alpha \approx 3$ , which is also known as the cubic law of returns. As Ibragimov et al. (2015) noted, the tail behavior of returns proposed by Gabaix et al. (2003) is essentially a characteristic of markets with well-developed mechanisms, where such heavy tails are generated by the trading of large market participants with size distributions that obey Zipf's law (a power law with  $\alpha \approx 1$ ).

This is an interesting characterization of the general power law behavior of the absolute returns of the exchange rate series under analysis. In particular, the results in table 2, show that in the complete sample, 43% of the series display an  $\alpha \in (2, 4)$ ; for subperiod I, there is a reduction to only 27%, and in subperiod II, the percentage of series increases to 46%. The increase observed in subperiod II may be a consequence of many countries becoming more active in the foreign exchange markets since the financial crisis. However, the exchange rates dynamics observed for several EM, particularly in subperiod II, is also the result of the adoption of a more flexible and relaxed foreign exchange system with respect to the USD. The resulting system represented a significant measure in safeguarding external competitiveness and preventing negative external shocks, promoting export growth, limiting the oscillation of exchange rates, and increasing foreign exchange reserves (Kohler, 2010).

Furthermore, over recent years, the shifting of capital flows to and from EM and the widening of currency mismatches also led some countries to reexamine foreign exchange market intervention strategies. In recent years, the focus of interventions has been to contain exchange rate volatility and provide insurance against exchange rate risk rather than to achieve a particular exchange rate. Hence, greater attention has been devoted to financial stability, as, for instance, in the case in Russia & Turkey (Mohanty & Berger, 2013). The results in table 2 confirm the statistically significant improvement in the tail index estimates from subperiod I to sub-period II for RUB and TRY.

The results in table 2 also identify series with greater heavy-tailedness ( $\alpha \leq 2$ ) in EM, which, as indicated in Ibragimov et al. (2015), may represent a challenge to the application of statistical and econometric methods. These results are to a certain extent not surprising given that several EM are heavily outward oriented, typically less diversified than more mature markets, and have less developed institutions. Moreover, since the cubic law of returns proposed by Gabaix et al. (2003) follows from a power law of vol-

umes ( $V_t$ ) with  $\alpha = 1.5^8$  and the square-root form-price impact  $r_t = kV_t^{1/2}$ , the differences observed for the tail indices in EM and DM countries may also result from the fact that the proposed square-root form-price impact may not be adequate for EM and may require adjustment. Hence, further analysis is needed to better understand the functional form of the cubic law of (exchange rate) returns (Gabaix et al., 2006; Ibragimov et al., 2015).

Additionally, further research on the implications of the structural changes detected in the tail behavior of the unconditional distribution of some of the exchange rate series (see section VC) from both a statistical and a policy perspective is necessary. For instance, applications of extreme value theory or conditional variance models typically depend on the stationarity assumption of the unconditional tail. However, a nonconstant  $\alpha$  implies a violation of covariance stationarity, which invalidates standard statistical inference based on regression analysis. Moreover, quantifying the tail index correctly is also of importance to risk managers and financial regulators.

## VI. Conclusion

In this paper, we introduce a new regression-based approach for the estimation of the tail index of heavy-tailed distributions; we derive its asymptotic properties and illustrate its good finite sample performance. We show that the proposed method of estimating  $\alpha$  presents four appealing features: (a) it does not involve order statistics; (b) it provides a bias reduction over the regression-based method proposed by Gabaix and Ibragimov (2012) and, a fortiori, over other regression-based estimators; (c) it is relatively robust to the choice of the subsample used to estimate  $\alpha$ ; and (d) when the effect of the slowly varying part in the Pareto-type model vanishes slowly (the so-called second-order behavior of the Taylor expansion), it continues to perform satisfactorily, whereas the Hill estimator rapidly deteriorates.

Moreover, given the novelty and flexibility of the procedure, two concrete avenues for future research involving this estimator can be explored: a feasible GLS estimator for  $\alpha$  and a different scheme to generate the regressor  $x_i$ . This suggests further generalizations to improve the performance of the tail index estimator when the effects of the slowly varying part in the Pareto-type model vanish slowly. (For preliminary results and further details on these extensions, see Nicolau & Rodrigues, 2015.)

We have also analyzed the effects of dependence and conditional heteroskedasticity on the properties of the proposed estimator through Monte Carlo simulations and suggest a variance estimator of the tail index estimator in this context. However, given the relevance and importance of this topic, further investigation is required. To illustrate the potential usefulness of the estimator, we also provide an empirical

<sup>8</sup>This power law for the volume corresponds to large heavy-tailedness, which suggests that trading is very concentrated.

analysis of the tail index of 74 absolute exchange rates returns series covering DM and EM. Overall, the empirical section presents a greater worldwide coverage of exchange rate series and a larger number of EM (59) than previous work. The results obtained corroborate existing evidence on the heavy-tailedness property differences between DM and EM countries, as well as within the EM group, and presents evidence of breaks in the tail index of some of the exchange rate returns series.

## REFERENCES

- Andrews, D. W. K., and J. C. Monahan, "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica* 60 (1992), 953–966.
- Beirlant, J., Y. Goegebeur, J. Segers, J. Teugels, D. Waal, and C. Ferro, *Statistics of Extremes: Theory and Applications* (Hoboken, NJ: John Wiley, 2004).
- Beirlant, J., C. Bouquiaux, and B. Werker, "Semiparametric Lower Bounds for Tail Index Estimation," *Journal of Statistical Planning and Inference* 136:3 (2006), 705–729.
- Beirlant, J., P. Vynckier, and J. L. Teugels, "Tail Index Estimation, Pareto Quantile Plots, and Regression Diagnostics," *Journal of the American Statistical Association* 91 (1996), 1659–1667.
- Danielsson, J., L. de Haan, L. Peng, and C. G. de Vries, "Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation," *Journal of Multivariate Analysis* 76 (2001), 226–248.
- Drees, H., L. de Haan, and S. Resnick, "How to Make a Hill Plot," *Annals of Statistics* 28 (2000), 254–274.
- Ebrechts, P., C. Klüppelberg, and T. Mikosch, *Modeling Extremal Events for Insurance and Finance* (Berlin: Springer, 2012).
- Gabaix, X., "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics* 114 (1999), 739–767.
- , "Power Laws in Economics and Finance," *Annual Review of Economics* 1 (2009), 255–293.
- Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley, "A Theory of Power Law Distributions in Financial Market Fluctuations," *Nature* 423 (2003), 267–270.
- Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley, "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics* 121 (2006), 461–504.
- Gabaix, X., and R. Ibragimov, "Log(Rank-1/2): A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business Economics and Statistics* 29:1 (2012), 24–39.
- Gabaix, X., and Y. M. Ioannides, "The Evolution of City Size Distributions," in V. Handerson and J. F. Thisse, eds., *Handbook of Urban and Regional Economics* (Amsterdam: Elsevier, North-Holland, 2004).
- Gabaix, X., and M. Maggiori, "International Liquidity and Exchange Rate Dynamics," *Quarterly Journal of Economics* 130:3 (2015), 1369–1420.
- Hall, P., "On Some Simple Estimates of an Exponent of Regular Variation," *Journal of the Royal Statistical Association (Series B)* 44 (1982), 37–42.
- Hartmann, P., S. Straetmans, and C. G. de Vries, "Heavy Tails and Currency Crises," *Journal of Empirical Finance* 17:2 (2010), 241–254.
- Hill, B. M., "A Simple General Approach to Inference about the Tail of a Distribution," *Annals of Statistics* 3 (1975), 1163–1174.
- Hill, J. B., "On Tail Index Estimation for Dependent, Heterogeneous Data," *Econometric Theory* 26 (2010), 1398–1436.
- Hols, M. C. A., and C. G. de Vries, "The Limiting Distribution of Extremal Exchange Rate Returns," *Journal of Applied Econometrics* 6:3 (1991), 287–302.
- Huisman, R., K. G. Koedijk, C. J. M. Kool, and F. Palm, "Tail-Index Estimates in Small Samples," *Journal of Business and Economic Statistics* 19 (2001), 208–216.
- Ibragimov, M., Z. Davidova, and R. Khamidov, *Heavy-Tailedness and Volatility in Emerging Foreign Exchange Markets: Theory and Empirics* (Kiev: EERC, 2010).
- Ibragimov, M., R. Ibragimov, and P. Kattuman, "Emerging Markets and Heavy Tails," *Journal of Banking and Finance* 37:7 (2013), 2546–2559.
- Ibragimov, M., R. Ibragimov, and R. Khamidov, "Measuring Inequality in CIS Countries: Theory and Empirics," GDN, working paper 54 (2011).
- Ibragimov, M., R. Ibragimov, and J. Walden, *Heavy-Tailed Distributions and Robustness in Economics and Finance* (Berlin: Springer, 2015).
- Jansen, D., and C. G. de Vries, "On the frequency of Large Stock Returns: Putting Booms and Busts into Perspective," this REVIEW 73 (1991), 18–24.
- Kearns, P., and A. Pagan, "Estimating the Density Tail Index for Financial Time Series," this REVIEW 79 (1997), 171–175.
- Koedijk, K. G., P. A. Stork, and C. G. de Vries, "Differences between Foreign Exchange Rate Regimes: The View from the Tails," *Journal of International Money and Finance* 11 (1992), 462–473.
- Kohler, M., "Exchange Rates during Financial Crises," *BIS Quarterly Review* (March 2010), 39–50.
- Kratz, M., and S. I. Resnick, "The qq-Estimator and Heavy Tails," *Stochastic Models* 12 (1996), 699–724.
- Loretan, M., and P. C. B. Phillips, "Testing the Covariance Stationarity of Heavy-Tailed Time Series," *Journal of Empirical Finance* 1 (1994), 211–248.
- Mason, D., "Laws of Large Numbers for Sums of Extreme Values," *Annals of Probability* 10 (1982), 754–764.
- Mikosch, T., and C. Stărică, "Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process," *Annals of Statistics* 28 (2000), 1427–1451.
- Mohanty, M., and B. Berger, "Central Bank Views on Foreign Exchange Intervention," Bank for International Settlements paper 73 (2013).
- Nguyen, T., and G. Samorodnitsky, "Tail Inference: Where Does the Tail Begin?" *Extremes* 15:4 (2012), 437–461.
- Nicolau, J., and P. M. M. Rodrigues, "A New Regression-Based Tail Index Estimator: An Application to Exchange Rates," Bank of Portugal working paper 14 (2015).
- Reinhart, C. M., and K. S. Rogoff, "The Modern History of Exchange Rate Arrangements: A Reinterpretation," *Quarterly Journal of Economics* 119:1 (2004), 1–48.
- Resnick, S., and C. Stărică, "Smoothing the Hill Estimator," *Applied Probability* 29 (1997), 271–293.
- Rosen, K. T., and M. Resnick, "The Size Distribution of Cities: An Explanation of the Pareto Law and Primacy," *Journal of Urban Economics* 8 (1980), 165–186.
- Samorodnitsky, G., and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (New York: Chapman and Hall, 1994).
- Stelzer, R., "On Markov-Switching ARMA Processes: Stationarity, Existence of Moments and Geometric Ergodicity," *Econometric Theory* 25:1 (2009), 43–62.
- Stoev, S. A., and G. Michailidis, "On the Estimation of the Heavy-Tail Exponent in Time Series Using the Max-Spectrum," *Applied Stochastic Models in Business and Industry* 26:3 (2010), 224–253.
- van der Vaart, A. W., and J. A. Wellner, *Weak Convergence and Empirical Processes with Applications to Statistics* (New York: Springer, 1996).