A NEW REGRESSION-BASED TAIL INDEX ESTIMATOR: AN APPLICATION TO EXCHANGE RATES

Working Papers 2015

João Nicolau | Paulo M. M. Rodrigues

BANCO DE PORTUGAL
EUROSYSTEM

14

14

A NEW REGRESSION-BASED TAIL INDEX ESTIMATOR: AN APPLICATION TO EXCHANGE RATES

Working Papers 2015

João Nicolau | Paulo M. M. Rodrigues

November 2015

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal
T +351 213 130 000 | estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisbon, 2015 • www.bportugal.pt

WORKING PAPERS | Lisbon 2015 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt •
Edition Economics and Research Department • ISBN 978-989-678-387-7 (online) • ISSN 2182-0422 (online)

A New Regression-Based Tail Index Estimator: An Application to Exchange Rates

João Nicolau
ISEG-Universidade de Lisboa
CEMAPRE

Paulo M. M. Rodrigues
Banco de Portugal
NOVA School of Business and Economics

November 2015

Abstract
In this paper, a new regression-based approach for the estimation of the tail index of heavy-tailed distributions is introduced. Comparatively to many procedures currently available in the literature, our method does not involve order statistics and can be applied in more general contexts than just Pareto. The procedure is in line with approaches used in experimental data analysis with fixed explanatory variables, and has several important features which are worth highlighting. First, it provides a bias reduction when compared to available regression-based methods and *a fortiori* over standard least-squares based estimators of the tail index. Second, it is more resilient to the choice of the tail length used in the estimation of the index than the widely used Hill estimator. Third, when the effect of the slowly varying function at infinity of the Pareto distribution (the so called second order behaviour of the Taylor expansion) vanishes slowly our estimator continues to perform satisfactorily, whereas the Hill estimator rapidly deteriorates. Fourth, our estimator performs well under dependence of unknown form. For inference purposes, we also provide a way to compute the asymptotic variance of the proposed estimator under time dependence and conditional heteroscedasticity. An empirical application of the procedure to exchange rates is also provided

JEL: C16, C18, C58

Acknowledgements: Acknowledgements: We thank 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, Uwe Hassler, Robin Lumsdaine, Morten Nielsen, Wing Wah Tham, Carlos Velasco and Chen Zhou for useful comments and suggestions. This work was partially funded by Fundação para a Ciência e Tecnologia through the project UID/Multi/00491/2013.

E-mail: nicolau@iseg.ulisboa.pt; pmrodrigues@bportugal.pt

1. Introduction

Over the last four decades there has been considerable interest in the estimation of the tail index of heavy-tailed distributions (which, henceforth, will be denoted by α). The interest results from the many areas of application of heavy-tailed distributions. For instance, these have been successfully used in computer science and telecommunications (Adler, Feldman, and Taqqu 1998; Resnick 1997; Chen et al. 2002), in finance and economics (Adler et al., 1998; Jansen and de Vries 1991), and in insurance (see Adler et al. 1998). This widespread interest has led over the years to the refinement and development of a number of tail index estimators, see, *inter alia*, the contributions of Hill (1975), Csörgo, Deheuvels, and Mason (1985), Kratz and Resnick (1996), Beirlant, Vynckier, and Teugels (1996), Feuerverger and Hall (1999), Crovella and Taqqu (1999) and Gabaix and Ibragimov (2012). For reviews of these methods see, *inter alia*, Haan *et al.* (2000), de Sousa and Michailidis (2004), Embrechts, Klüppelberg and 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 infer an appropriate value for α from the properties of the resulting graph (Kratz and Resnick, 1996 and Beirlant et al., 1996). However, although many of these estimation methods exhibit interesting asymptotic properties (such as, *e.g.*, consistency), their finite sample performance is questionable (see, for instance, the results in Kearns and Pagan, 1997, and Huisman, Koedijk, Kool and Palm, 2001).

One of the most popular and frequently used approaches in the empirical literature is the Hill estimator (Hill, 1975). The nice theoretical properties of this estimator - consistency (Deheuvels, Haeusler, and Mason 1988; and Mason 1982) and asymptotic normality (Hall 1982) - have led researchers (see, e.g. Drees, de Haan, and Resnick 2000; and Resnick and Stărică, 1997) to develop improved variants and to show that these modifications work well in the Pareto case. However, available tail index estimation procedures present problems when applied to data drawn from distributions other than the Pareto. Hence, the correct identification of α remains a challenging and empirically relevant quest. In this paper, we introduce a new regression-based procedure which overcomes several empirical difficulties of currently available tail index estimators.

The new method introduced in this paper to estimate α does not involve order statistics, and can be applied in more general contexts than just Pareto. There are three important features of our method which can be summarized as follows: first, it provides a bias reduction when compared to the regression-based method proposed by Gabaix and Ibragimov (2012), and *a fortiori* over other regression based estimators; second, it is more resilient to the choice of the subsample used to estimate α than the widely used Hill estimator; in other words, the mean square error of our estimator is less sensitive to the incorrect determination of the tail length than the Hill estimator and to some extent than the estimator of Gabaix and Ibragimov (2012); and third, when the effect of the slowly varying function at infinity of the Pareto distribution vanishes slowly (the so called second order behavior of the Taylor expansion) our estimator continues to perform satisfactorily, whereas the Hill estimator rapidly deteriorates.

The remainder of the paper is organized as follows. In section 2 we briefly present three widely applied tail index estimators, which will be used later as benchmarks for comparison with the new procedure proposed in this paper. Section 3 introduces the new tail index estimator and discusses its asymptotic properties; Section 4 provides a detailed Monte Carlo analysis of the finite sample bias of the tail index estimators discussed in Sections 2 and 3, as well as an analysis of the impact of empirically relevant features frequently found in economic and financial time series, such as time dependence and conditional heteroscedasticity; Section 5 discusses possible directions for further generalizations of the procedure developed in this paper; and Section 6 illustrates the potential usefulness of the approach in an empirical application on daily exchange rate returns series from 21 countries considering the USD as numeraire. Finally, Section 7 summarizes the main results. A technical appendix collects the proofs of the results put forward throughout the paper.

2. Tail Index Estimators

Although there is a vast literature on tail index estimators (see, *e.g.*, Beirlant *et al.*, 2004, chapter 4 for an overview), in this section we briefly describe three of the most popular procedures used in empirical work, and which will be used as benchmarks for comparison with the new approach that will be introduced in section 3.

In what follows, a heavy-tailed distribution is defined as a distribution function F such that $1-F$ is regularly varying at infinity with index $-\alpha$, i.e,

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

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

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

$$\widehat{\alpha}_H := \left(\frac{1}{m} \sum_{j=1}^{m} \log X_{(j)} - \log X_{(m+1)} \right)^{-1} \tag{2}$$

where $m := [\kappa n]$, with $\kappa \in (0,1)$, is the number of highest order statistics used in the estimation of α. Hall (1982) showed that for $m/n \to 0$ as m, $n \to \infty$ that $\sqrt{m}\left(\frac{\widehat{\alpha}_H}{\alpha} - 1\right)$ is asymptotically normal distributed; and Hill (2010) provides results on the asymptotic properties of the Hill estimator for heavy-tailed heterogenous dependent processes. Moreover, available evidence suggests that the Hill estimator is most effective when the underlying distribution is either Pareto or close to Pareto (Drees, Haan and Resnick, 2000). However, if the distribution is as in (1) the Hill estimator is only approximately a maximum likelihood estimator and its accuracy will become less clear.

Remark 2.1: The hill estimator measures the average increase of the Pareto quantile plot above a certain threshold and can be interpreted as a slope estimator of the linear part of the Pareto quantile plot.

Remark 2.2: The adequate choice of m has been an important topic of research and several approaches for its determination in the i.i.d. context have been put forward; see, e.g., Danielsson, Haan, Peng and de Vries (2001) and Nguyen and Samorodnitsky (2012). This is an important concern since the convergence in distribution of the Hill statistic given in (2) critically hinges on the rate at which the nuisance parameter m grows with the total sample size.

Remark 2.3: A further shortcoming of this approach is the use of order statistics as these require sorting the data which may become computationally expensive (it requires at least $O(n \log n)$ steps), and destroys the tie ordering of the data and their temporal structure (Stoev and Michailidis, 2008).

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 which need to be specified for the adequate performance of the procedure.

As a consequence, 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, *e.g.,* Rosen and Resnick, 1980, and Gabaix, 1999), *i.e.*,

$$\log(t-\gamma) = \varrho - \alpha \log X_{(t)} + error_t, \tag{3}$$

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 $\log \frac{t}{n} \approx \log(C) - \alpha \log(X_{(t)})$, $t = 1, ..., N$; see Gabaix and Ibragimov (2012). This regression has been popular, for instance, in the urban literature for the analysis of Zipf's law (Zipf, 1949). The statistical properties of the OLS estimators of (3) have been analyzed in Gabaix and Ioannides (2004), Nishiyama, Osada and Sato (2008) and Gabaix and Ibragimov (2012). For reference purposes, in what follows, we define the OLS estimator of α computed from (3) as $\widehat{\alpha}_{\gamma=0}$.

Finally, the third procedure we consider is a recent important contribution by Gabaix and Ibragimov (2012) who propose an improved version of regression (3), which consists in the estimation of (3) with the optimal shift of $\gamma = 1/2$. The motivation for the proposal of this improved regression results from the fact that although the OLS estimator $\widehat{\alpha}_{\gamma=0}$ computed from (3) with $\gamma = 0$ is consistent, it suffers from important 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 (3) significantly reduces this bias, while maintaining the good asymptotic properties of the estimator (see also the Monte Carlo results provided in Section 4 below). For comparison purposes we define the OLS estimator of α computed from (3) with $\gamma = 1/2$ as $\widehat{\alpha}_{\gamma=1/2}$.

Note that one common feature to the approaches in (2) and (3) is that all of these methods rely on order statistics computed from i.i.d. random variables drawn from Pareto type distributions.

3. The New Tail Index Estimator

3.1. The Estimator

To introduce the new tail index estimator we propose in this paper consider first the simple case of a sequence of i.i.d. random variables $\{X_t;\ t = 1, 2, ..., n\}$ drawn from a Pareto-type tail, such as,

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

Applying logarithms to (4) and rearranging we obtain that

$$\log \bar{F}_n(x_i) = \alpha \log x_0 - \alpha \log x_i + \left[\log \bar{F}_n(x_i) - \log \bar{F}(x_i)\right], \qquad i = 1, 2, ... \tag{5}$$

where $\bar{F}_n(x_i) := \frac{1}{n}\sum_{t=1}^{n} \mathcal{I}_{\{X_t > x_i\}}$ and $\mathcal{I}_{\{.\}}$ 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 $\left[\log \bar{F}_n(x_i) - \log \bar{F}(x_i)\right] =: \varepsilon_i$ is the error term. Standard methods to estimate α in a regression frameworks such as (3) consist in treating x_i as an order statistics, say $X_{(i)}$, and as a result, $\bar{F}_n\left(X_{(i)}\right) = i/n$. 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 will propose next treats x_i as a nonrandom variable, which considerably simplifies the analysis of the statistical properties of the resulting estimator, as we will see below.

To implement our approach, we need to generate x_i, which is used to compute $\log \bar{F}_n(x_i)$, according to a deterministic scheme. A method that one can consider for generating the deterministic values of x_i is, for example, $x_i = x_0 + i\Delta$, $i = 0, 1, ..., m$, so that x_i varies between x_0 and $x_0 + m\Delta$ (assuming that x_0 is known). In this case, x_i is defined as an additive sequence with a fixed increment of Δ. A regression model such as suggested in (5) is valid under mild conditions such as, $n, m \to \infty$ and $\Delta \to 0$ and leads to consistent OLS estimators of α. Notice that m and Δ are under our control and we may generate as many x_i values as we like. However, the choice of x_i as just described is not optimal.

Alternatively and as with Importance Sampling methods, a variance reduction technique used in Monte Carlo integration, instead of spreading the sample points of x_i out evenly we may concentrate the distribution of x_i in parts of the state space of X which are of most "importance".

Since x_i is such that $x_i := F^{-1}(u_i)$, where F^{-1} is the quantile function of F and $u_i \in (0, 1)$, assuming that F is a Pareto distribution as in (4) we have that,

$$x_i := F^{-1}(u_i) = (1 - u_i)^{-\frac{1}{\alpha}} x_0.$$

Note that we avoid treating u_i as an uniformly distributed random variable to keep x_i as a deterministic realization. Thus, we consider

$$u_i := \frac{i}{m}, \qquad i = 1, 2, ..., m - 1.$$

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, (5) can be written as

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

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

However, equation (6), is infeasible given the dependence of z_i on α. Thus, to make (6) feasible, we deal with α in z_i as a nuisance parameter and treat z_i as a generated regressor (Pagan, 1984). In other words, the implementation of our procedure consist of two-steps: In the first step, α is computed from a consistent estimator (such as, *e.g.*, the Hill estimator or as indicated below in remark 3.1), and we denote the resulting estimate as $\tilde{\alpha}$, which is used to generate the regressor $\tilde{z}_i = \tilde{\alpha}^{-1} \log(1 - u_i)$. The second step then corresponds to the actual OLS estimation of α from a feasible version of (6), *i.e.*, from,

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

We will denote the resulting estimator from (7) as $\widehat{\alpha}_{Pareto}$. Note that the substitution of α by $\tilde{\alpha}$ computed in the first step has little impact on the estimation of α from (7), since its effect vanishes asymptotically as long as $\tilde{\alpha} \xrightarrow{p} \alpha$ (see the proofs of Theorems 1 and 2 in the Appendix). When X is governed by a strict Pareto distribution, x_0 is the left endpoint of the support of X and can be estimated as $\hat{x}_0 = \min(X_1, ..., X_n)$ and optimal properties are achieved, *i.e.*, the bias rapidly converges to zero and the variance of $\hat{\alpha}$ is of order $1/n$ (just as under maximum likelihood estimation).

One further interesting aspect of our methodology is that z_i, and to some extent even $\hat{z}_i$, can be treated as fixed explanatory variables. Another point worth mentioning is that the whole probabilistic structure of the model can be derived. In other words, considering the $(m-1) \times 1$ vector of errors $\varepsilon := (\varepsilon_1, ..., \varepsilon_{m-1})'$, we can show that $\sqrt{n}\varepsilon \xrightarrow{d} N(0, \mathbf{\Sigma})$, where the variance/covariance matrix $\mathbf{\Sigma}$ is known exactly (no estimation is required) regardless of whether x_0 is known or unknown (see Lemma 1 below). This is an important result as it allows us *i*) to obtain the $\lim Var(\hat{\alpha})$, and *ii*) to conduct generalized least squares estimation if required.

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

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

where $a > 0$, $\beta > \alpha$ and $b \in \mathbb{R}$, which includes, among many others, the non-trivial alpha-stable distribution and the Student-t distribution. The parameters b and β govern the second order behavior of the Taylor expansion and aim to reflect the deviation from strict Pareto tail behavior.

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

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

where as before $y_i := \log \bar{F}_n(x_i)$, $x_i := (1-u_i)^{-\frac{1}{\alpha}} x_0$, and $z_i := \alpha^{-1}\log(1-u_i)$. The most important difference between (7) and (9) is that the error term $\tilde{\varepsilon}_i$ in the latter case is,

$$\tilde{\varepsilon}_i := \varepsilon_i + \eta_i, \tag{10}$$

where $\varepsilon_i := \log \bar{F}_n(x_i) - \log \bar{F}(x_i)$ and $\eta_i := \log\left(1 + b(1-u_i)^{\frac{\beta}{\alpha}} x_0^{-\beta} + o\left(x_i^{-\beta}\right)\right)$. We show that the additional term, η_i, in (10) is responsible for finite sample bias of our estimator, but vanishes asymptotically under appropriate conditions (see the proof of Theorem 2). We will show in Section 4, that this finite sample bias is smaller than that of the Hill estimator and of the regression-based methods considered in Section 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 used to generate $y_i := \log \bar{F}_n(x_i)$, $i = 1, 2, ..., m-1$. We estimate x_0 as $\hat{x}_0 := \hat{F}^{-1}(1-\kappa)$, $0 < \kappa < 1$, *i.e.* $\hat{x}_0$ is the empirical quantile of order $1-\kappa$. The smaller the value of κ the better the approximation of (8) to the tail of the Pareto law will be and, consequently, the closer $\tilde{\varepsilon}_i$ in (10) will be to ε_i. In other words, as $\kappa \to 0$ we have that $\hat{x}_0 \to \infty$ and $\tilde{\varepsilon}_i \to \varepsilon_i$. In section 5 we provide further insights on the choice of κ.

3.2. Asymptotic Properties

3.2.1. Limits under the i.i.d. case. To characterize the asymptotic properties of our new estimator we need first to consider the results provided next in Lemmas 1 and 2.

LEMMA 1. *Let $\{X_t;\ t = 1, 2, ..., n\}$ be a sequence of Pareto distributed i.i.d. random variables. Then, as $n \to \infty$:*

(i)

$$\sqrt{n}\begin{pmatrix} \bar{F}_n(x_1) - \bar{F}(x_1) \\ \bar{F}_n(x_2) - \bar{F}(x_2) \\ \vdots \\ \bar{F}_n(x_{m-1}) - \bar{F}(x_{m-1}) \end{pmatrix} \xrightarrow{d} N(0, \mathbf{A}) \tag{11}$$

where $\mathbf{A} := [a_{ij}]_{(m-1)\times(m-1)}$, $a_{ij} := F(x_i \wedge x_i)\bar{F}(x_j)$, *and* $x_i \wedge x_j := \min\{x_i, x_j\}$.

(ii)

$$\sqrt{n}\varepsilon \xrightarrow{d} N(0, \mathbf{\Sigma}) \tag{12}$$

where $\varepsilon :=(\varepsilon_1, ..., \varepsilon_{m-1})'$, $\varepsilon_i := \log \bar{F}_n(x_i) - \log \bar{F}(x_i)$ *and* $\mathbf{\Sigma} := [\sigma_{ij}]_{(m-1)\times(m-1)}$. *If* x_0 *is known, the parameters* σ_{ij} *are given by*

$$\sigma_{ij} := \frac{q}{m-q}, \qquad q := i \wedge j,\ i,j = 1,2,...,m-1. \tag{13}$$

Otherwise, if x_0 *is estimated as* $\hat{x}_0 := F^{-1}(1-k)$,

$$\sigma_{ij} := \frac{k(q-m)+m}{k(m-q)}, \qquad q := i \wedge j,\ i,j = 1,2,...,m-1. \tag{14}$$

(iii) Considering $\hat{x}_0 := F^{-1}(1-k)$, *for large* n, *it follows that,*

$$\lim Var\left(\sum_{i=1}^{m-1}\sqrt{n}\varepsilon_i\right) = \frac{-k+m\left(2k-2-\sum_{i=1}^{m-1}\frac{1}{i}\right)+m^2(2-k)}{k} + o(1) = O\left(\frac{m^2}{k}\right). \tag{15}$$

LEMMA 2. *Considering* $w_i := \begin{pmatrix} 1 & z_i \end{pmatrix}'$, *with* $z_i := \alpha^{-1}\log(1-i/m)$ *it follows that:*

(i)

$$\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i' = \begin{bmatrix} 1 & -\frac{1}{\alpha} \\ -\frac{1}{\alpha} & \frac{2}{\alpha^2} \end{bmatrix} + o(1). \tag{16}$$

(ii) Let $\{X_t; t = 1,2,..,n\}$ *be a sequence of Pareto distributed i.i.d. random variables, then*

$$\lim Var\left(\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right) = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{17}$$

where

$$a := \sum_{i=1}^{m-1} i + \sum_{j=1}^{m-2}\sum_{i=1}^{j}\frac{i}{m-i} = m^2 + o\left(m^2\right),$$

$$b := \sum_{i=1}^{m-1} z_i i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{i}{m-i} = -\frac{2m^2}{\alpha} + o\left(m^2\right),$$

$$c := \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j} z_i\frac{i}{m-i} + \sum_{j=1}^{m-1} z_j\frac{j}{m-j}\sum_{i=j}^{m-1} z_i = \frac{5m^2}{\alpha^2} + o\left(m^2\right).$$

(iii) Let $\{X_t; t = 1, 2, .., n\}$ be a sequence of i.i.d. random variables with distribution given by (8). Thus, for x large enough we have that,

$$\lim Var\left(\sum_{i=1}^{m-1} w_i \sqrt{n}\varepsilon_i\right) = \frac{m}{k}\begin{bmatrix} d & e \\ e & f \end{bmatrix} \tag{18}$$

where

$$\begin{aligned}
d &:= m - 1 + \sum_{j=1}^{m-2}\sum_{i=1}^{j}\frac{1}{m-i} = 2m + o(m), \\
e &:= \sum_{i=1}^{m-1} z_i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{1}{m-i} = -\frac{2m}{\alpha} + o(m), \\
f &:= \frac{z_1}{m-1}\sum_{i=1}^{m-1} z_i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{z_i}{m-i} + \sum_{j=1}^{m-1}\frac{z_j}{m-j}\sum_{i=j}^{m-1} z_i = \frac{6m}{\alpha^2} + o(m).
\end{aligned}$$

Hence, the results in Lemmas 1 and 2 allow us to state the following two theorems with the properties of the new estimator.

THEOREM 1. *Let $\{X_t; t = 1, 2, .., n\}$ be a sequence of Pareto distributed i.i.d. random variables and assume that $\tilde{\alpha} \xrightarrow{p} \alpha$. It follows, as $n \to \infty$, that the OLS estimator computed from (7) is consistent and normally distributed, i.e.,* ***(i)*** *$\hat{\alpha}_{Pareto} \xrightarrow{p} \alpha$ and* ***(ii)*** *$\sqrt{n}\left(\hat{\alpha}_{Pareto} - \alpha\right) \xrightarrow{d} N\left(0, 2\alpha^2\right)$.*

THEOREM 2. *Let $\{X_t; t = 1, 2, .., n\}$ be a sequence of i.i.d. random variables with survival function given by (8). Furthermore, assuming that $n = cm^{\gamma}$, with $\gamma > 1$ ($k = m/n \to 0$). It follows as $n \to \infty$ that the OLS estimator computed from (9) is consistent and normally distributed, i.e.,* ***(i)*** *$\hat{\alpha}_{Pareto} \xrightarrow{p} \alpha$ and* ***(ii)*** *$\sqrt{m}\left(\hat{\alpha}_{Pareto} - \alpha\right) \xrightarrow{d} N\left(0, 2\alpha^2\right)$.*

Remark 3.1: It follows from the proofs of Theorems 1 and 2 (see Appendix) that it is not necessary to assume that $\tilde{\alpha} \xrightarrow{p} \alpha$. In fact, any value $\tilde{\alpha} > 0$ leads to a consistent $\hat{\alpha}_{Pareto}$ estimator. Of course, there may be a cost of efficiency when $\tilde{\alpha}$ is an arbitrary value. For this reason our estimator may be obtained either as we have indicated in the previous section, or using the following procedure: First, set $\tilde{\alpha} = c > 0$, where c is any positive value, say $c = 4$, for example. Use $\tilde{\alpha}$ to generate the regressor $\tilde{z}_i = \tilde{\alpha}^{-1} \log(1 - u_i)$ and run regression (7) to obtain $\tilde{\alpha} = \tilde{\alpha}_{Pareto}$ and a new sequence of $\tilde{z}_i$. Run again regression (7) to obtain the final estimate of $\tilde{\alpha}_{Pareto}$. Some simulations carried out by the authors have confirmed that this procedure ensures rapid convergence, even when the starting value for $\tilde{\alpha}$ is nonsense (e.g. $\tilde{\alpha} = 200$). 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 help of any other estimator; and second, and most importantly, it simplifies the proof of results, as we do not need to specify additional conditions on $\tilde{\alpha}$.

Remark 3.2: Theorem 1 establishes that the variance of $\hat{\alpha}_{Pareto}$ is of order $O\left(n^{-1}\right)$, where n represents the complete sample size. This is a natural result given that all data are sampled from an exact Pareto distribution. In a more general setting, when the tail behaviour is of the Pareto-Levy form, the variance of the $\hat{\alpha}_{Pareto}$ estimator is of order $O\left(m^{-1}\right)$, where $m = [\kappa n]$ and $\kappa \in (0, 1)$, according to Theorem 2.

Remark 3.3: Interestingly, the proposed estimator reaches the same asymptotic variance as the $\widehat{\alpha}_{\gamma=1/2}$ estimator proposed by Gabaix and Ibragimov (2012), with an important methodological difference: our estimator was obtained assuming a Pareto-Levy tail, whereas in Gabaix and Ibragimov (2012) only a strict power law decline was considered (i.e. imposing $b = 0$, $a = 1$ and $o(x^{\beta}) = 0$, in equation 8).

Remark 3.4: 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, for instance, Beirlant et al. (2004).

A practical issue is how to define m. Although m can be set as large or as small as desired, our results show that there is no particular advantage in setting m arbitrarily large when compared to the sample size n (notice that if n is small and m is large the procedure will lead to many ties of y_t - different values of x_i leading to the same value of y_t). Thus, we set $m := [\kappa n]$ where

$0 < \kappa < 1$ and $[.]$ corresponds to the largest integer of the argument; see section 5 for more details. To the best of our knowledge no theoretical results have been proposed yet for the dependent case.

3.2.2. Dependent Data. The results in Theorems 1 and 2, derived in the i.i.d. context, are of importance since the proposed estimator is new in the literature and some of its basic properties needed to be established first 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 $\{X_t\}$ may exhibit dependence of unknown form.

We performed an extensive Monte Carlo analysis (available upon request) to address two specific questions : i) how does dependence affect the asymptotic variance of $\hat{\alpha}_{Pareto}$, and ii) how does dependence affect the precision of $\hat{\alpha}_{Pareto}$ and the Hill estimator. Our main conclusions are the following: i) The dependence in the data, via AR or GARCH dynamics, has important impacts on the limiting distributions of $\hat{\alpha}_H$ 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 κ increase. This impact vanishes completely when the original process is replaced by the residuals from an AR model. The limiting distribution of $\hat{\alpha}_H$ is more affected, especially when κ is small, but like the $\hat{\alpha}_{Pareto}$ estimator, the autocorrelation effect tends to decline as the sample size and κ increase. On the contrary, the GARCH effect has a strong impact on the asymptotic variance of both estimators (for results on the Hill estimator see e.g. Quintos, Fan and Phillips, 2001, and Hill, 2010), especially when κ is in the range of "optimal values". Some of our conclusions are in line with the results of Kearns and Pagan (1997). ii) The dependence in the data via AR or GARCH has an impact on the optimal choice of κ: the higher the dependence the higher is the optimal value for κ. The change in the optimal choice of κ is even more marked in the IGARCH case. Interestingly, dependence in the data does not seem to affect the quality of the $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$ estimates as long as κ is properly adjusted. Another point is that the proposed estimator performs better than the Hill estimator for almost all values of κ.

To sum up, our findings point to a valid estimator $\hat{\alpha}_{Pareto}$ of the tail index, but to an inconsistent estimator for the variance of $\hat{\alpha}_{Pareto}$ when the i.i.d.

hypothesis is wrongly assumed in the presence of dependence (as we expected). This is a problem that is also common to the other estimators described in Section 2. Hence, in what follows we develop asymptotic results for the variance of $\hat{\alpha}_{Pareto}$ under general conditions.

A careful analysis of the proofs of Theorems 1 and 2 shows that the random quantity $\frac{1}{m}\sum_{i=1}^{m-1} w_i \varepsilon_i$, 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. Hence, some of the conditions we impose 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 σ-fields generated by the random variables $X_a, ..., X_b$, and define

$$a(n) := \sup\left\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^{\infty}, k \geq 1\right\}.$$

Thus, the process $\{X_t\}$ is strongly mixing if $a(n) \to 0$.

Theorem 3. *Let $\{X_t\}$ be a strictly stationary process with distribution F and survival function given by (8) and assume that $\{X_t\}$ satisfies the strong mixing condition $a(n) \to 0$. Furthermore, assume that $n = cm^{\gamma}$, with $\gamma > 1$ ($k = m/n \to 0$).* ***(i)*** *It follows as $n \to \infty$ that the OLS estimator computed from (9) is consistent i.e. $\hat{\alpha}_{Pareto} \xrightarrow{p} \alpha$.* ***(ii)*** *In addition, if $a(n) = o\left(n^{-6-\varepsilon}\right)$, $\varepsilon \in (0,1)$ then the estimator is normally distributed, i.e.*

$$\sqrt{m}\left(\hat{\alpha}_{Pareto} - \hat{\alpha}\right) \xrightarrow{d} N\left(0, \mathcal{V}_{22}\right)$$

where $\mathcal{V}_{22}$ is the element $(2,2)$ of

$$\mathcal{V} := \Xi \mathcal{J} \Xi,$$

with $\Xi := \begin{bmatrix} 2 & \alpha \\ \alpha & \alpha^2 \end{bmatrix}$, $\mathcal{J} := \lim_{m\to\infty} J_m$; $J_m := \frac{1}{m}\Sigma_{s=1}^{m-1}\Sigma_{i=1}^{m-1} \lim E\left(V_s V_i'\right)$, *and* $V_i = w_i \varepsilon_i$.

Remark 3.5: We note that α-mixing is the weakest among the most frequently used mixing conditions and that the α-mixing coefficients in part (ii) of Theorem 3 decrease at an arithmetic rate. Therefore, the conditions we require in Theorem 3, are broad enough to include most processes with interest in economics and finance. For example, under some weak assumptions ARMA, GARCH, and Markov-Switching processes, among many others, can easily satisfy the conditions of Theorem 3 (see, for instance, Fan and Yao, 2005; Boussama, 1998; and Stelzer, 2009)[1].

To estimate J_m in the presence of autocorrelation and heteroscedasticity of unknown form we consider a HAC estimator as introduced by Andrews and Monahan (1992). It is based on the following procedure. Consider that $\hat{\theta} = \begin{pmatrix} \hat{\vartheta} & \hat{\alpha} \end{pmatrix}'$ is a $\sqrt{m}$-consistent estimator of θ. First, one estimates a bth order VAR model for $V_t = w_t \hat{\varepsilon}_t$ (2×1 vector), $V_t = \sum_{i=1}^{b} \hat{A}_i V_{t-i} + V_t^*$ for $t = b+1, ..., m$, where V_t^* is the corresponding residual vector. Second, one computes a standard kernel-based HAC estimator, say $\hat{J}_m^* \left(\hat{S}_m \right)$, based on the VAR residual vector V_t^*, *i.e.*,

$$\hat{J}_m^* \left(\hat{S}_m \right) = \frac{m}{m-2} \sum_{i=-m+1}^{m-2} \mathcal{K} \left(\frac{j}{\hat{S}_m} \right) \hat{\Gamma}_j^*$$

where

$$\hat{\Gamma}_j^* = \begin{cases} \frac{1}{m} \sum_{t=j+1}^{m-1} V_t^* V_{t-j}^{*\prime} & \text{for } j \geq 0 \\ \frac{1}{m} \sum_{t=-j+1}^{m-1} V_t^* V_{t-j}^{*\prime} & \text{for } j < 0 \end{cases},$$

$\mathcal{K}(\cdot)$ is a real-valued kernel, defined as

$$\mathcal{K}(x) = \frac{25}{12\pi^2 x^2} \left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right)$$

and $\hat{S}_m$ is a data-dependent bandwidth parameter, defined as $\hat{S}_m := 1.3221 (\hat{c} m)^{1/5}$, where

$$\hat{c} = \frac{4\hat{\rho}^2 \hat{\sigma}^4}{(1-\hat{\rho})^8} \bigg/ \frac{\hat{\sigma}^4}{(1-\hat{\rho})^4},$$

and (ρ, σ^2) denote the autoregressive and innovation variance parameters of the second component of V_t, i.e. $\{z_i \hat{\varepsilon}_i\}$. Third, one *recolors* the estimator $\hat{J}_m^* \left(\hat{S}_m \right)$ to obtain the VAR prewhitened kernel estimator of J_m :

$$\hat{J}_m \left(\hat{S}_m \right) = \hat{D} \hat{J}_m^* \left(\hat{S}_m \right) \hat{D}', \qquad \hat{D} = \left(I_2 - \sum_{i=1}^{b} \hat{A}_i \right)^{-1}.$$

1. In some of these cases, it is possible to prove that the processes are beta mixing with geometric rate which, in turn, implies strong mixing with arithmetic rate.

THEOREM 4. *Considering that* $\sqrt{m}\left(\hat{A}_i - A_i\right) = O(1)$, $i = 1,...,n$, $\left(I_2 - \sum_{i=1}^{b} A_i\right)$ *is nonsingular and the conditions of Theorem 3 hold, then* $\hat{J}_m\left(\hat{S}_m\right) - J_m \xrightarrow{p} \mathbf{0}$.

4. Monte Carlo Analysis

In this section, we perform a Monte Carlo study to assess and compare the root mean square error (RMSE) of the four estimators discussed in the previous sections, as a function of κ. In particular, we compare the standard Hill estimator, $\widehat{\alpha}_H$; the OLS estimator, $\widehat{\alpha}_{\gamma=0}$, computed from the log-log rank-size regression, $\log t = \varrho - \alpha \log X_{(t)} + error_t$; the OLS estimator, $\widehat{\alpha}_{\gamma=1/2}$, computed from the log (Rank-1/2) regression, $\log(t - 1/2) = \varrho - \alpha \log X_{(t)} + error_t$; and finally the new estimator introduced in this paper, $\widehat{\alpha}_{Pareto}$, computed from (9).

We select several heavy tailed distributions which satisfy equation (8) as DGPs, namely, the Student-t, the alpha-stable, and the Burr distribution. These distributions are frequently used in the extreme-value literature (see Beirlant *et al.*, 2004). To generate data from a stable distribution with index α we use the method of Samorodnitsky and Taqqu (1994), i.e.

$$X_t = \frac{\sin(\alpha\theta_t)}{\left[\cos(\theta_t)\right]^{1/\alpha}}\left(\frac{\cos\left[(1-\alpha)\theta_t\right]}{z_t}\right)^{(1-\alpha)/\alpha} \tag{19}$$

where θ_t is uniform on $(-\pi/2, \pi/2)$ and z_t is an exponential variate with mean 1. The Monte Carlo experiments performed in this section consist of the following steps: 1) Generate a sample of size n, with $n \in \{500, 2000, 5000\}$, from a heavy tailed distribution; 2) Fix a value for κ in the set

$$\begin{aligned}\kappa &= \{0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.125, 0.15, , \\ &\quad 0.175, 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375, 0.4\}. \end{aligned} \tag{20}$$

Notice that κ 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; 3) Estimate α from a subsample of size $m = [\kappa n]$ using the four estimators, $\widehat{\alpha}_H$, $\widehat{\alpha}_{\gamma=0}$, $\widehat{\alpha}_{\gamma=1/2}$ and $\widehat{\alpha}_{Pareto}$; 4) Using the same distribution (and the same parameters) repeat steps 1) and 3) 1000 times and calculate the RMSE associated with the estimation of α, $RMSE(\hat{\alpha}_i) = \sqrt{\sum_{t=1}^{1000}(\hat{\alpha}_{i,t} - \alpha)^2 / 1000}$ where i refers to one of the four estimators under analysis; 5) Select another value for κ in the above set and repeat all previous steps; 6) Select another DGP and/or another value of α and repeat all previous steps.

First, we compare the performance of the $\widehat{\alpha}_H$, $\widehat{\alpha}_{\gamma=0}$ and $\widehat{\alpha}_{\gamma=1/2}$ estimators. Figures 1 - 3 present the results obtained for these estimators.

[Please insert Figures 1 - 3 about here]

The main conclusions that we can draw from these Figures are the following: (i) the minimum RMSE of the Hill estimator, $\widehat{\alpha}_H$, and of the log rank-size regression, $\widehat{\alpha}_{\gamma=1/2}$, computed from (3) with $\gamma = 1/2$ are approximately equal, although they are generally attained at different values of κ; (ii) in contrast to the regression-based estimators, the RMSE of the Hill estimator rapidly deteriorated as κ moves away from its minimum; and as a consequence, the estimators $\widehat{\alpha}_{\gamma=0}$ and $\widehat{\alpha}_{\gamma=1/2}$ are more resilient to variations in κ; (iii) the estimator $\widehat{\alpha}_{\gamma=1/2}$ displays better performance than $\widehat{\alpha}_{\gamma=0}$ for almost all values of κ. There is significant bias reduction when the estimator $\widehat{\alpha}_{\gamma=1/2}$ is used comparatively to $\widehat{\alpha}_{\gamma=0}$, as was also indicated by Gabaix and Ibragimov (2012). Thus, these results show that the OLS estimator of α computed from $\log(t-1/2) = \varrho - \alpha \log X_{(t)} + error_t$ ($\widehat{\alpha}_{\gamma=1/2}$) is in general the best amongst the three estimators we have just analyzed.

Given the dominant performance of $\widehat{\alpha}_{\gamma=1/2}$, in what follows we compare this estimator to the new estimator, $\widehat{\alpha}_{Pareto}$, introduced in this paper, through the ratio $RMSE\left(\widehat{\alpha}_{\gamma=1/2}\right)/RMSE\left(\widehat{\alpha}_{Pareto}\right)$ under various scenarios for κ, sample sizes and DGPs. Figures 4 - 6 illustrate the results obtained.[2]

[Please insert Figures 4 - 6 about here]

These figures show that the new estimator introduced performs generally better than $\widehat{\alpha}_{\gamma=1/2}$. Note that in general, for the distributions considered $\frac{RMSE\left(\widehat{\alpha}_{\gamma=1/2}\right)}{RMSE\left(\widehat{\alpha}_{Pareto}\right)} > 1$.

2. In Figures 9 to 11 a comparison of the RMSE of the Hill estimator ($\widehat{\alpha}_H$), and the estimator introduced in this paper ($\widehat{\alpha}_{Pareto}$) are also provided.

5. Future Developments

In this section, we briefly discuss two potential directions of generalization of the tail index estimator proposed in this paper, which will be addressed in future research. One relates to the use of generalized least squares (GLS) in the estimation of the tail index, and the other to the generation of x_i from distributions other than the Pareto.

5.1. GLS Estimation

The first is related to Lemma 1, which establishes that $\sqrt{n}\tilde{\varepsilon} \sim N(0, \mathbf{\Sigma})$ where $\mathbf{\Sigma} := [\sigma_{ij}]_{(m-1)\times(m-1)}$ and σ_{ij} are as given in Lemma 1 (ii). Knowing exactly the elements of the matrix $\mathbf{\Sigma}$ has the advantage of allowing us to consider generalized least squares (GLS) estimation which leads to more efficient estimators of α than OLS, but not necessarily to estimators with smaller mean square errors. The GLS estimator is

$$\begin{pmatrix} \hat{\vartheta}_{gls} \\ \hat{\alpha}_{gls} \end{pmatrix} = \left(\tilde{W}'\mathbf{\Sigma}^{-1}\tilde{W}\right)^{-1}\tilde{W}\mathbf{\Sigma}^{-1}y$$

where $\tilde{W}$ is an $(m-1)\times 2$ matrix with rows $\begin{pmatrix} 1 & \widetilde{z}_i \end{pmatrix}$, $i = 1, 2, ..., m-1$. Unreported preliminary results suggest that while the asymptotic variance of the GLS estimator decreases (as expected) comparatively to that of the OLS estimator, the mean square error may not necessarily decrease given the impact of the asymptotic bias. This topic requires further in depth investigation.

5.2. The generating distribution of x_i

The second development focuses on the generation of x_i. In section 3 we suggested generating x_i according to the rule $x_i := (1-u_i)^{-\frac{1}{\alpha}} x_0$, where $u_i := i/m$, $i = 1, 2, ..., m-1$. The idea was that if X has a Pareto distribution then the optimal choice to sample x_i is precisely $x_i := F^{-1}(u_i)$ where F is the Pareto distribution function and F^{-1} the corresponding quantile function. However, the generation of x_i may be generalized to other distribution functions. To illustrate the potential of this generalization, we generate x_i considering a Student t-distribution. For comparative purposes denote this estimator as $\widehat{\alpha}_{Student}$. We performed a Monte Carlo analysis similar to the one presented in the previous section, considering a sample of $n = 1000$. The data was generated according to a t-distribution with $\alpha = 1, 2, 3, 4$ degrees of freedom and the proposed estimator described in section 3, $\widehat{\alpha}_{Pareto}$, compared to $\widehat{\alpha}_{Student}$ using again the ratio $\frac{RMSE(\widehat{\alpha}_{Pareto})}{RMSE(\widehat{\alpha}_{Student})}$. Figure 12 summarizes the results.

[Please insert Figure 12 about here]

We observe from this figure that $\widehat{\alpha}_{Student}$ does in general perform better than $\widehat{\alpha}_{Pareto}$, particularly when $\kappa > 0.1$. These results can be partially justified by the fact that the DGP and the distribution function considered to generate $x_i := F^{-1}(u_i)$ for the computation of $\widehat{\alpha}_{Student}$ coincide, which rarely occurs in practice. Nevertheless, the main idea of this limited Monte Carlo study is to illustrate the potential and simplicity of generalization of the methodology presented in Section 3. Hence, this opens new avenues to improve the performance of the tail index estimator when the effects of the slowly varying part in the Pareto type model vanish slowly. This issue requires further investigation, in particular it will be interesting to obtain results as in Theorem 2 and Lemma 1 for the same cases but under different distributions.

6. Empirical Application

To illustrate the empirical performance of the estimator introduced earlier, we provide an empirical application to exchange rate returns. Exchange rates have been widely analyzed in the literature. In the context of studies on heavy-tails, several important contributions have been made; see, among others, Hols and de Vries (1991), Koedijk, Stork and de Vries (1992), Loretan and Phillips (1994), Cotter (2005), Ibragimov, Davidova and Khamidov (2010), Hartmann, Straetmans and de Vries (2009), and Ibragimov, Ibragimov and Kattuman (2013). The tail index is of importance as it can be used as a measure of an economy's vulnerability to shocks, *i.e.*, the likelihood of extreme movements and changes occurring.

In this section, we analyze the tail properties of daily exchange rate returns series of 21 countries considering the US dollar (USD) as base currency[3]. The currencies considered may be classified into three groups: i) **Developed markets**: Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Danish krone (DKK), Euro (EUR), Hong Kong dollar (HKD), Great Britain pound (GBP), Japanese yen (YEN), New Zealand (NZD), Norwegian kroner (NOK), Singapore dollar (SGD) and Swedish krona (SEK); ii) **Emerging markets:** Brasil real (BRL), Chilean peso (CLP), Colombian peso (COP), Mexican peso (MXP), Polish zloty (PLN), South Korean won (KRW), South Africa (ZAR); and iii) **Frontier markets:** Argentine peso (ARS) and Ukrainian hryvna (UAH).[4]

In Figures 13 and 14 we present several exchange rates to USD and corresponding returns series, respectively, which illustrate the different behavior of these series over the sample under analysis.

[Please insert Figures 13 and 14]

The period of analysis is from December 31, 1993 to February 13, 2015 and all data is obtained from Datastream. Table 1 presents some descriptive statistics for the series under analysis.

[Please insert Table 1 about here]

The tail indices of the exchange rate growth series in developed, emerging and frontier markets are estimated, using the four estimators discussed in the text ($\widehat{\alpha}_H$, $\widehat{\alpha}_{Pareto}$, $\widehat{\alpha}_{\gamma=0}$ and $\widehat{\alpha}_{\gamma=1/2}$), based on a 5%, 10%, 15% and 20% truncation level (κ) for the extreme observations. However, for presentation purposes and given the results of the Monte Carlo simulations provided in the previous section, we will only discuss results for $\widehat{\alpha}_H$ with $\kappa = 0.1$, and for $\widehat{\alpha}_{Pareto}$ with $\kappa = 0.2$. The choice of these more generous truncation levels results from the Monte Carlo observation that the minimum MSE for these estimators, in the case of volatility in the data (as is the case for the daily exchange rate

3. Results with other base currencies (EUR, YEN, GBP and CHF) can be obtained from the authors.

4. This classification is based on Morgan Stanley's Markets classification http://www.msci.com/products/indexes/market_classification.html.Note however that care needs to be taken with these classification given that over the sample period considered some of the countries may not have always belonged to the group in which they are presently included.

returns series considered in this section), is obtained with a higher truncation level. Tables 2 and 3 present the tail index estimates.

[Please insert Tables 2 and 3 about here]

From Tables 2 and 3 we observe that the right tail indices of the exchange rate growth series for the 21 countries analyzed for the complete sample and for the three subperiods 1994-1999, 2000-2007 and 2008-2015, show considerable difference in the three groups of countries under analysis (developed markets ($\mathcal{DM}$), emerging markets ($\mathcal{EM}$) and frontier markets ($\mathcal{FM}$)).[5]

For the complete sample (1994 - 2015) we note that the HKD, COP, KRW, ARS, BRL, MXP and UAH display tail index estimates such that $\widehat{\alpha} < 2$. Moreover, it is also interesting to observe from these results that the currencies which present smaller tail indices are the ones linked to the USD.[6]

However, since currency crises have always been a characteristic of the international monetary system, considering a more detailed analysis of the tail index by subsamples may reveal further insights. Hence, given our sample size and the dramatic episodes of the Latin American Tequila Crisis following Mexico's peso devaluation in 1994-95, the Asian financial crisis in 1997-98, the Russian financial crisis of 1998 and, more recently, the global financial crisis in 2008-09 which forced sharp depreciations in many advanced as well as developing economies, we will reestimate the tail indices over three subsamples: 1) 1994 - 1999; 2) 2000 - 2007; and 3) 2008 - 2015.

Subsample 1 - 1994 to 1999

The 90s were a particularly turbulent period. The subsample from 1994 to 1999 includes the Mexican crisis in 1995 (which spread to other economies in the region, affecting particularly Argentina); and the Asian and Russian financial crises in 1997-98, which impacted Brazil in 1998-99. The consequences of these crises were so severe that they originated changes in the macroeconomic policies of countries in these regions, especially in terms of their exchange rate policies (Frankel, Fajnzylber, Schmukler and Serven, 2001, and Frenkel and

5. Note that results for the left tail can be obtained from the authors as the conclusions are qualitatively similar to the ones based on the right tail.

6. Note that we classify an exchange rate to have $\widehat{\alpha} < 2$, when both $\widehat{\alpha}_H < 2$ and $\widehat{\alpha}_{Pareto} < 2$.

Rapetti, 2011). For instance, after the 1994-95 crisis, Mexico changed to a floating exchange rate regime (Frenkel and Rapetti, 2011), and a similar policy was followed in 1999, by Brazil, Colombia and Chile.

Over this period our results show (Tables 2 - 3) that currencies with an $\widehat{\alpha} < 2$, are HKD, SGD, BRL, ZAR, KRW, MXP, PLN, ARS, COP and UAH. As expected, over this period the currencies with the smallest tail index (i.e., with the potentially highest risk profile) belonged to Asian countries, Latin American countries and the Ukrain.

Subsample 2 - 2000 to 2007

This period 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. However, from 2002 until the financial disruption in 2008, developed and developing countries went through a prosperous period without crises.

This is reflected in the results of Tables 2-3. For this period our results indicate that currencies with an $\widehat{\alpha} < 2$, were only HKD, COP, ARS and UKH (note that over this period 1 out of the 6 currencies in the $\mathcal{EM}$ group display $\widehat{\alpha} < 2$). It becomes clear from these results that in general, comparatively to the previous period, the tail index has increased suggesting a possible risk reduction in most currencies.

Subsample 3 - 2008 - 2015

Finally, the last subsample includes the recent financial crisis. Interestingly, financial contagion following the collapse of Lehman Brothers was short and by 2009 many developing countries had recovered access to the international financial system at low interest rates. This was likely the consequence of the switch made by the countries belonging to the $\mathcal{EM}$ group to flexible managed floating regimes (see e.g. Reinhart and Rogoff, 2004 (updated country chronologies), for worldwide details on exchange rate regimes) and due to the accumulation of foreign exchange reserves over this period.

Our results in Tables 2 and 3 suggest that in this period the currencies with an $\widehat{\alpha} < 2$ are HKD, COP, KRW, ARS and UAH (again a much smaller number than in subsample 1). Comparatively to the results obtained for subsample 1 we observe that the number of currencies with $\widehat{\alpha} < 2$ is smaller in this period (in specific, we observe a reduction from 10 to 5 currencies with $\widehat{\alpha} < 2$).

A final comment relates to the similarities or otherwise of the left and right tail indices. In foreign exchange markets, there seems to exist a consensus that volatility is symmetric with respect to positive and negative shocks (*e.g.* Bollerslev and Domowitz, 1993, and Anderson *et al.*, 2001). This is probably justified by the "two-sided nature of the foreign exchange market" *i.e.*, by the observation that for bilateral exchange rates positive returns for one currency are necessarily negative returns for the other. This is probably also the reason why one could expect the right and left tail indices not to differ.

However, as recently pointed out by Wang and Yang (2009), this may not necessarily be always the case, given that despite the bilateral nature of exchange rates, at least two reasons which may justify the presence of asymmetry in bilateral exchange rates can be suggested. The first is the greater economic importance of some currencies over others; and the second are central bank interventions. Many studies report that central bank interventions lead to higher volatility, particularly when there is a depreciation in the domestic currency and not when there is an appreciation. Since central banks intervene on one side of the market, but not on the other, interventions may lead to an asymmetric relationship between exchange rate return and volatility.

We also observe that for several currencies the right tail appears to be slightly more heavy-tailed than the left tail. This type of asymmetry is the opposite of the asymmetric behavior typically found in financial markets and may indicate regulatory interventions in these currency markets (as suggested by *e.g.* Ibragimov *et al.* 2013), however, this feature requires further investigation.

7. Conclusion

In this paper a new regression-based approach for the estimation of the tail index of heavy-tailed distributions is introduced, its asymptotic properties are derived and its good finite sample performance illustrated.

We show that the proposed method to estimate α presents the following features: First, it does not involve order statistics. Second, it provides a bias reduction over the regression-based method proposed by Gabaix and Ibragimov (2012), and *a fortiori* over other regression based estimators. Third, it is relatively robust to the choice of the subsample used to estimate α. Fourth, 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) our estimator continues to perform satisfactorily, whereas the Hill estimator rapidly deteriorates.

Moreover, given the novelty and flexibility of the procedure, we also propose two concrete avenues for future research involving this estimator, namely, a feasible GLS estimator for α, and a different scheme to generated the regressor x_i. We have analyzed the effects of dependence and conditional heteroscedasticity on the properties of the proposed estimator, through Monte Carlo simulations. However, given the relevance and importance of this topic further investigation is required.

To illustrate the potential of the estimator we provide an empirical application in which we analyze the tail index of 21 exchange rates returns series.

References

Adler, R. J., R. E. Feldman and M.S. Taqqu (1998), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Boston, MA: Birkhauser.

Andersen, T. G., T. Bollerslev, F. X. Diebold and P. Labys (2001), The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association* 96, 42-55.

Beirlant, J., P. Vynckier and J. L. Teugels, (1996), Tail Index Estimation, Pareto Quantile Plots, and Regression Diagnostics. *Journal of the American Statistical Association*, 91, 1659–1667.

Beirlant, J., Y. Goegebeur, J. Segers, J. Teugels, D. Waal and C. Ferro, (2004), *Statistics of Extremes: Theory and Applications*. Wiley Series in Probability and Statistics. John Wiley & Sons.

Bollerslev, T. and I. Domowitz (1993) Trading Patterns and Prices in the Interbank Foreign Exchange Market. *Journal of Finance* 48, 1421-1443.

Chen, Q., H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger (2002), The Origin of Power-Laws in Internet Topology Revisited. *Proceedings of IEEE Infocom* 2002.

Crovella, M. E. and M. S. Taqqu, (1999), Estimating the Heavy Tail Index from Scaling Properties. *Methodology and Computing in Applied Probability* 1, 55–79.

Csörgõ, S., P. Deheuvels and D. Mason, (1985), Kernel estimates of the tail index of a distribution. *Annals of Statistics* 13(3), 1050–1077.

Danielsson, J., L. de Haan, L. Peng, and C. G. de Vries, (2001), Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation. *Journal of of Multivariate Analysis*, 76, 226–248.

Davis R., Mikosch T. (2009) Extreme Value Theory of GARCH processes. In T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.): *Handbook of Financial Time Series*, 187-200. New York: Springer.

Deheuvels, P., E. Haeusler, and D. M. Mason, (1988), Almost Sure Convergence of the Hill Estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*, 104, 371–381.

Drees, H. and E. Kaufmann, (1998), Selecting the Optimal Sample Fraction in Univariate Extreme Value Estimation. *Stochastic Processes and Their Applications*, 75, 149–172.

Drees, H., L. de Haan and S. Resnick, (2000), How to make a Hill plot. *The Annals Statistics* 28, 254–274.

Embrechts, P., C. Klüppelberg and T. Mikosch (2012), *Modeling extremal events for insurance and finance*, Springer.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. New York: Wiley.

Feuerverger, A., and P. Hall (1999), Estimating a Tail Exponent by Modelling Departure from a Pareto Distribution. *The Annals of Statistics* 27, 760–781.

Frankel, J.A., E. Fajnzylber, S.L. Schmukler and L. Serven (2001), Verifying exchange rate regimes. *Journal of Development Economics* 66, 351 - 386.

Frenkel, R. and Rapetti, M. (2011), Exchange Rate Regimes in Latin America, in José Antonio Ocampo and Jaime Ros (eds.): *Handbook of Latin American Economics*, Oxford University Press.

Gabaix, X. (1999), Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics* 114, 739–767.

Gabaix, X. and Y.M. Ioannides (2004), The evolution of city size distributions, in: V. Handerson, J.F. Thisse (Eds.), in: *Handbook of Urbanand Regional Economics*, vol. 4, Elsevier, North Holland, Chap. 53.

Gabaix, X. and R. Ibragimov (2012), Log(Rank-1/2): A Simple Way to Improve the OLS Estimation of Tail Exponents. *Journal of Business Economics and Statistics 29(1), 24 - 39.*

de Haan, L., H. Drees, and S. Resnick (2000), How to make a Hill plot. *The Annals of Statistics* 28(1), 254–274.

Hall, P. (1982), On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Association* (Series B) 44, 37–42.

Hardy, G. H. and E. M. Wright (1979), *An Introduction to the Theory of Numbers*. Oxford: Oxford University Press.

Hartmann, P., S. Straetmans and C.G. De Vries (2010) Heavy tails and currency crises. *Journal of Empirical Finance* 17(2), 241-254.

Hill, B. M. (1975), A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* 3, 1163–1174.

Hill, J.B. (2010), On tail index estimation for dependent, heterogeneous data. *Econometric Theory* 26, 2010, 1398–1436.

Hill, J.B. (2014), Tail Index Estimation for a Filtered Dependent Time Series. *Statistica Sinica* 25, 609-630.

Hols, M.C.A. and C.G. de Vries (1991), The limiting distribution of extremal exchange rate returns, Journal of Applied Econometrics 6(3), 287-302.

Hsieh, P.-H. (1999), Robustness of Tail Index Estimation. *Journal of Computational and Graphical Statistics* 8, 318–332.

Hsing, T. (1991). On tail index estimation using dependent data. Annals of Statistics 19, 1547-1569.

Huisman, R., K.G. Koedijk, C.J.M. Kool, F. Palm (2001), Tail-index estimates in small samples. *Journal of Business and Economic Statistics* 19, 208–216.

Ibragimov M., Z. Davidova and R. Khamidov (2010), Heavy-Tailedness and Volatility in Emerging Foreign Exchange Markets: Theory and Empirics. - Kiev: EERC, 51p.

Ibragimov, M., R. Ibragimov, P. Kattuman, (2013), Emerging markets and heavy tails. *Journal of Banking & Finance* 37(7), 2546–2559.

Jansen, D., and C. de Vries (1991), On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective. *Review of Economics and Statistics* 73, 18-24.

Kearns, P. and A. Pagan (1997), Estimating the Density Tail Index for Financial Time Series. *Review of Economics and Statistics* 79, 171-175.

Koedijk, K. G., P. A. Stork and C. G. de Vries, (1992), Differences between foreign exchange rate regimes: the view from the tails. *Journal of International Money and Finance* 11, 462–473.

Kosorok, M.R. (2008), *Introduction to Empirical Processes and Semiparametric Inference*, Springer, New York.

Kratz, M. and S. I. Resnick (1996), The qq-estimator and heavy tails. *Stochastic Models* 12, 699-724.

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

Mason, D. (1982), Laws of Large Numbers for Sums of Extreme Values. *Annals of Probability* 10, 754–764.

Mandelbrot, B. B. (1960), The Pareto-Lévy law and the distribution of income. *International Economic Review* 1(2), 79–106.

Nguyen, T., and G. Samorodnitsky (2012), Tail Inference: where does the tail begin?. *Extremes* 15 (4): 437-461.

Nishiyama, Y., S. Osada, and Y. Sato (2008), OLS estimation and the t-test revisited in rank-size rule regression. *Journal of Regional Science* 48(4), 691–716.

Pagan, A. R. (1984), Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25, 221–247.

Patnaik, I., A. Shah, A. Sethy, and V. Balasubramaniam (2011), The exchange rate regime in Asia: From crisis to crisis. *International Review of Economics and Finance,* 32–43.

Reinhart, C.M. and K.S. Rogoff (2004), The Modern History of Exchange Rate Arrangements: A Reinterpretation. *Quarterly Journal of Economics* 119(1), 1-48.

Resnick, S. (1997), Heavy Tail Modeling and Teletraffic Data. *The Annals of Statistics* 25, 1805–1869.

Resnick, S. and C. Stărică (1997), Smoothing the Hill Estimator. *Applied Probability* 29, 271–293.

Rose, A. K. (2011), Exchange Rate Regimes in the Modern Era: Fixed, Floating, and Flaky. *Journal of Economic Literature* 49(3), 652–672.

Rosen, K.T. and M. Resnick (1980), The Size distribution of Cities: An Explanation of the Pareto Low and Primacy, *Journal of Urban Economics* 8, 165-186.

Samorodnitsky, G. and M.S. Taqqu (1994) *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance*. Chapman and Hall, New York - London.

de Sousa, B. and G. Michailidis (2004), A diagnostic plot for estimating the tail index of a distribution. *Journal of Computational and Graphical Statistics* 13(4), 974–995.

Stoev, S. and G. Michailidis (2009), On the estimation of the heavy-tail exponent in time series using the max-spectrum. *Applied Stochastic Models in Business and Industry*.

Wang, J. and M. Yang (2009), Asymmetric volatility in the foreign exchange markets. *Journal of International Financial Markets, Institutions & Money* 19, 597 - 615.

Wellner, J.A. (1977), A Glivenko-Cantelli Theorem and Strong Laws of Large Numbers for Functions of Order Statistics. *The Annals of Statistics* 5 (3), 473 - 480.

Wooldridge, J. M., (2010), *Econometric Analysis of Cross Section and Panel Data*, Second Edition, MIT Press.

Zipf, G. K. (1949), Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology, Reading, MA: Addison-Wesley

Appendix: Technical Appendix

Before presenting the proofs of the main results put forward in the paper it will be convenient to establish some preliminary Lemmas first.

LEMMA A.1. *The following results hold,*

$$(i)\ \sum_{i=1}^{m-1} \log i = -(m-1) + (m-1)\log(m-1) + \frac{1}{2}\log(m-1) + O(1); \quad (A.1)$$

$$(ii)\ \sum_{i=1}^{m-1} (\log i)^2 = (m-1)\log^2(m-1) - 2(m-1)\log(m-1) + 2(m-1) + \frac{\log^2(m-1)}{2} + O(1). (A.2)$$

Proof of Lemma A.1

(i) The result in (A.1) follows directly from Stirling's formula (see Feller, 1968, p.52), *i.e.*, since,

$$\sum_{i=1}^{m-1} \log i = \log((m-1)!)$$

and

$$(m-1)! \sim e^{-(m-1)}(m-1)^{(m-1)+1/2}\sqrt{2\pi}$$

it follows that,

$$\log((m-1)!) - \log(e^{-(m-1)}(m-1)^{(m-1)+1/2}\sqrt{2\pi}) \to 0$$

and therefore,

$$\begin{aligned}\sum_{i=1}^{m-1} \log i &= -(m-1) + (m-1)\log(m-1) + \log\left(\sqrt{2\pi}\right) + \frac{1}{2}\log(m-1) + o(1) \\ &= -(m-1) + (m-1)\log(m-1) + \frac{1}{2}\log(m-1) + O(1).\end{aligned}$$

(ii) Regarding (A.2), applying the Euler-Maclaurin formula for asymptotic expansions of sums to $\sum_{i=1}^{m-1} \log^2 i$, we immediately obtain that,

$$\sum_{i=1}^{m-1} \log^2 i = (m-1)\log^2(m-1) - 2(m-1)\log(m-1) + 2(m-1) + \frac{\log^2(m-1)}{2} + O(1).$$

■

LEMMA A.2. *Considering $u_i := i/m$ and m a positive integer, we establish that,*

$$
\begin{aligned}
(i) \quad \sum_{i=1}^{m-1} \log(1-u_i) &= -m + O(\log m); && \text{(A.3)}\\
(ii) \sum_{i=1}^{m-1} (\log(1-u_i))^2 &= 2m + O\left(\log^2 m\right); && \text{(A.4)}\\
(iii) \quad \sum_{i=1}^{m-1} \frac{u_i}{1-u_i} &= 1 - m + m\sum_{k=1}^{m-1}\frac{1}{k}. && \text{(A.5)}
\end{aligned}
$$

Proof of Lemma A.2

(i) To prove (A.3) note that,

$$\sum_{i=1}^{m-1} \log(1-u_i) = -(m-1)\log m + \sum_{i=1}^{m-1} \log i.$$

Thus, from (A.1) it follows that,

$$
\begin{aligned}
\sum_{i=1}^{m-1} \log(1-u_i) &= -(m-1)\log m - (m-1) + (m-1)\log(m-1) + \frac{1}{2}\log(m-1) + O(1)\\
&= -(m-1) - (m-1)[\log m - \log(m-1)] + \frac{1}{2}\log(m-1) + O(1)\\
&= -m + \left(1 + \frac{1}{2}\log(m-1)\right) - (m-1)[\log m - \log(m-1)] + O(1)\\
&= -m + O(\log(m)).
\end{aligned}
$$

(ii) For the proof of (A.4) we establish that,

$$
\begin{aligned}
\sum_{i=1}^{m-1} (\log(1-u_i))^2 &= \sum_{i=1}^{m-1} (\log(m-i) - \log m)^2\\
&= \sum_{i=1}^{m-1} \log^2(m-i) - 2\log m \sum_{i=1}^{m-1} \log(m-i) + (m-1)\log^2 m\\
&= \sum_{i=1}^{m-1} \log^2 i - 2\log m \sum_{i=1}^{m-1} \log i + (m-1)\log^2 m.
\end{aligned}
$$

Thus, from (A.1) and (A.2) it follows that,

$$\begin{aligned}
\sum_{i=1}^{m-1} (\log(1-u_i))^2 &= \sum_{i=1}^{m-1} \log^2 i - 2\log m \left[-(m-1) + (m-1)\log(m-1) + \frac{1}{2}\log(m-1)\right] \\
&\quad + (m-1)\log^2 m + O(1) \\
&= 2(m-1) + \frac{1}{2}\left[\log^2(m-1) - 2\log m \log(m-1)\right] \\
&\quad + 2(m-1)\left[\log m - \log(m-1)\right] \\
&\quad + (m-1)\left[\log^2 m + \log^2(m-1) - 2\log m \log(m-1)\right] + O(1) \\
&= 2m + O(\log^2 m).
\end{aligned}$$

(iii) Finally, regarding (A.5) note that,

$$\begin{aligned}
\sum_{i=1}^{m-1} \frac{u_i}{1-u_i} &= \sum_{i=1}^{m-1} \frac{i}{m-i} \\
&= \sum_{k=1}^{m-1} \sum_{i=1}^{m-k} \frac{1}{i} \\
&= 1 - m + m \sum_{k=1}^{m-1} \frac{1}{k}.
\end{aligned}$$

Considering that the harmonic sum

$$\begin{aligned}
\sum_{k=1}^{m-1} \frac{1}{k} &= \log(m-1) + \mathcal{C} + \frac{1}{2(m-1)} - \frac{1}{12(m-1)^2} + \frac{1}{120(m-1)^4} + \ldots \\
&= \log(m-1) + \mathcal{C} + o(1)
\end{aligned}$$

where $\mathcal{C} \approx 0.5772156649$ is the Euler–Mascheroni constant (see e.g., Hardy and Wright, 1978), it follows that,

$$\sum_{i=1}^{m-1} \frac{u_i}{1-u_i} = (1-m) + m\log(m-1) + m\mathcal{C} + o(1).$$

■

Proof of Lemma 1

(i) According to Donsker's theorem it follows that,

$$\sqrt{n}\begin{pmatrix} \bar{F}_n(x_1)-\bar{F}(x_1) \\ \bar{F}_n(x_2)-\bar{F}(x_2) \\ \vdots \\ \bar{F}_n(x_{m-1})-\bar{F}(x_{m-1}) \end{pmatrix} = -\sqrt{n}\begin{pmatrix} F_n(x_1)-F(x_1) \\ F_n(x_2)-F(x_2) \\ \vdots \\ F_n(x_{m-1})-F(x_{m-1}) \end{pmatrix} \xrightarrow{d} N(0,\ \mathbf{A})$$

where

$$\mathbf{A}=\begin{pmatrix} F(x_1)\bar{F}(x_1) & F(x_1)\bar{F}(x_2) & F(x_1)\bar{F}(x_3) & \cdots & F(x_1)\bar{F}(x_{m-1}) \\ F(x_1)\bar{F}(x_2) & F(x_2)\bar{F}(x_2) & F(x_2)\bar{F}(x_3) & \cdots & F(x_2)\bar{F}(x_{m-1}) \\ \bar{F}(x_1)F(x_3) & F(x_2)\bar{F}(x_3) & F(x_3)\bar{F}(x_3) & \cdots & F(x_3)\bar{F}(x_{m-1}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F(x_1)\bar{F}(x_{m-1}) & F(x_2)\bar{F}(x_{m-1}) & F(x_3)\bar{F}(x_{m-1}) & \cdots & F(x_{m-1})\bar{F}(x_{m-1}) \end{pmatrix}$$

(see, for example, Kosorok, 2008, Chapter 2). Notice that $\mathbf{A}:=[a_{ij}]_{(m-1)\times(m-1)}$ and $a_{ij}:=\lim Cov\left(\sqrt{n}F_n(x_i),\sqrt{n}F_n(x_j)\right)=F(x_i\wedge x_j)-F(x_i)F(x_j)=F(x_i)\bar{F}(x_j),\ x_i\leq x_j$. Considering $\mathbf{g}(x):=(\log x_1,\log x_2,...,\log x_{m-1})'$ using the delta method for the multivariate case we establish that,

$$\sqrt{n}\left(\mathbf{g}\left(\bar{F}_n(x)\right)-\mathbf{g}\left(F(x)\right)\right)=\sqrt{n}\begin{pmatrix} \log\bar{F}_n(x_1)-\log\bar{F}(x_1) \\ \log\bar{F}_n(x_2)-\log\bar{F}(x_2) \\ \vdots \\ \log\bar{F}_n(x_{m-1})-\log\bar{F}(x_{m-1}) \end{pmatrix} \xrightarrow{d} N(0,\mathbf{\Sigma})$$

where

$$\mathbf{\Sigma}=\frac{\partial\mathbf{g}\left(\bar{F}_n(x)\right)}{\partial x'}\mathbf{A}\left(\frac{\partial\mathbf{g}\left(\bar{F}_n(x)\right)}{\partial x'}\right)'=\begin{bmatrix} \frac{F(x_1)}{\bar{F}(x1)} & \frac{F(x_1)}{\bar{F}(x_1)} & \cdots & \frac{F(x_1)}{\bar{F}(x_1)} \\ \frac{F(x_1)}{\bar{F}(x_1)} & \frac{F(x_2)}{\bar{F}(x_2)} & \cdots & \frac{F(x_2)}{\bar{F}(x_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{F(x_1)}{\bar{F}(x_1)} & \frac{F(x_2)}{\bar{F}(x_2)} & \cdots & \frac{F(x_{m-1})}{\bar{F}(x_{m-1})} \end{bmatrix}.$$

For x_0 known we have that $F(x_i)=1-\left(\frac{x_0}{x_i}\right)^{\alpha}=\frac{i}{m}$ and $\bar{F}(x_i)=1-\frac{i}{m}$, and therefore,

$$\frac{F(x_i)}{\bar{F}(x_i)}=\frac{i}{m-i}$$

and

$$\mathbf{\Sigma} = \begin{bmatrix} \frac{1}{m-1} & \frac{1}{m-1} & \frac{1}{m-1} & \cdots & \frac{1}{m-1} \\ \frac{1}{m-1} & \frac{2}{m-2} & \frac{2}{m-2} & \cdots & \frac{2}{m-2} \\ \frac{1}{m-1} & \frac{2}{m-2} & \frac{3}{m-3} & \cdots & \frac{3}{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m-1} & \frac{2}{m-2} & \frac{3}{m-3} & \cdots & \frac{m-1}{m-(m-1)} \end{bmatrix}.$$

(ii) For $\hat{x}_0 := F^{-1}(1-k) = k^{-1/\alpha} x_0$ we have, after some simplifications, that,

$$\begin{aligned} F(x_i) &= 1 - \left(\frac{x_0}{x_i}\right)^{\alpha} = 1 - \left(\frac{x_0}{(1-u_i)^{-\frac{1}{\alpha}} k^{-1/\alpha} x_0}\right)^{\alpha} = 1 - \left(1 - \frac{i}{m}\right) k \\ \bar{F}(x_i) &= \left(1 - \frac{i}{m}\right) k, \end{aligned}$$

so that

$$\frac{F(x_i)}{\bar{F}(x_i)} = \frac{k(i-m)+m}{k(m-i)},$$

and

$$\mathbf{\Sigma} = \begin{bmatrix} \frac{k(1-m)+m}{k(m-1)} & \frac{k(1-m)+m}{k(m-1)} & \cdots & \frac{k(1-m)+m}{k(m-1)} \\ \frac{k(1-m)+m}{k(m-1)} & \frac{k(2-m)+m}{k(m-2)} & \cdots & \frac{k(2-m)+m}{k(m-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{k(1-m)+m}{k(m-1)} & \frac{k(2-m)+m}{k(m-2)} & \cdots & \frac{k((m-1)-m)+m}{k(m-(m-1))} \end{bmatrix}.$$

(iii) Finally, to prove (15), note that,

$$\begin{aligned} Var\left(\sum_{i=1}^{m-1} \sqrt{n}\varepsilon_i\right) &= \sum_{i=1}^{m-1} Var\left(\sqrt{n}\varepsilon_i\right) + 2\sum_{i=1}^{m-1}\sum_{j=1}^{i-1} Cov\left(\sqrt{n}\varepsilon_i, \sqrt{n}\varepsilon_j\right) \\ &= \sum_{i=1}^{m-1} \sigma_{ii} + 2\sum_{i=1}^{m-1}\sum_{j=1}^{i-1} \sigma_{ij}. \qquad \text{(A.6)} \end{aligned}$$

Hence, (A.6) corresponds to the sum of all elements of $\mathbf{\Sigma}$. To simplify (A.6) note that $\sigma_{11} := \frac{k(1-m)+m}{k(m-1)}$ is repeated $(2m-3)$ times in $\mathbf{\Sigma}$, $\sigma_{22} := \frac{k(2-m)+m}{k(m-2)}$ is repeated $(2m-5)$ times, and so on. Therefore,

$$\begin{aligned} Var\left(\sum_{i=1}^{m-1} \sqrt{n}\varepsilon_i\right) &= \sum_{i=1}^{m-1} \frac{k(i-m)+m}{k(m-i)} (2m-(2i+1)) \\ &= \frac{-k + m\left(2k - 2 - \sum_{i=1}^{m-1} \frac{1}{i}\right) + m^2(2-k)}{k} = O\left(\frac{m^2}{k}\right). \end{aligned}$$

■

Proof of Lemma 2

(i) Based on the results in (i) and (ii) of Lemma A.2 it follows that,

$$\begin{aligned}\sum_{i=1}^{m-1} w_i w_i' &= \begin{bmatrix} m & \sum_{i=1}^{m-1} z_i \\ \sum_{i=1}^{m-1} z_i & \sum_{i=1}^{m-1} z_i^2 \end{bmatrix} \\ &= \begin{bmatrix} m & \frac{1}{\alpha}\sum_{i=1}^{m-1}\log(1-u_i) \\ \frac{1}{\alpha}\sum_{i=1}^{m-1}\log(1-u_i) & \frac{1}{\alpha^2}\sum_{i=1}^{m-1}(\log(1-u_i))^2 \end{bmatrix} \\ &= \begin{bmatrix} m & -\frac{1}{\hat{\alpha}}m \\ -\frac{1}{\hat{\alpha}}m & \frac{2m}{\hat{\alpha}^2} \end{bmatrix} + \begin{bmatrix} 0 & O(\log m) \\ O(\log m) & O(\log^2 m) \end{bmatrix}.\end{aligned}$$

(ii) Considering,

$$\mathbf{W}' := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_{m-1} \end{bmatrix},$$

we have that,

$$Var\left(\sum_{i=1}^{m-1} w_i \sqrt{n}\varepsilon_i\right) = \mathbf{W}'\mathbf{\Sigma}\mathbf{W}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_{m-1} \end{bmatrix} \begin{bmatrix} \frac{1}{m-1} & \frac{1}{m-1} & \cdots & \frac{1}{m-1} \\ \frac{1}{m-1} & \frac{2}{m-2} & \cdots & \frac{2}{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m-1} & \frac{2}{m-2} & \cdots & \frac{m-1}{m-(m-1)} \end{bmatrix} \begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_{m-1} \end{bmatrix}$$

$$= \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

where $a = \sum_{i=1}^{m-1} i + \sum_{j=1}^{m-2}\sum_{i=1}^{j}\frac{i}{m-i}$; $b = \sum_{i=1}^{m-1} z_i i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{i}{m-i}$; $c = \sum_{j=1}^{m-2}\sum_{i=1}^{j} z_i \frac{i}{m-i} + \sum_{j=1}^{m-1}\frac{j}{m-j}\sum_{i=j}^{m-1} z_i$; and $d = \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j} z_i\frac{i}{m-i} + \sum_{j=1}^{m-1} z_j \frac{j}{m-j}\sum_{i=j}^{m-1} z_i$.

We observe that a can be simplified as $a = m^2 - 1 - m\sum_{i=1}^{m-1} 1/i = m^2 + o(m^2)$. However, b and c do not have a closed-form, but can be determined numerically so that $b = -2m^2/\alpha + o(m^2)$ and $c = 5m^2/\alpha^2 + o(m^2)$.

(iii) We have

$$Var\left(\sum_{i=1}^{m-1} w_i \sqrt{n}\varepsilon_i\right) = \mathbf{W}'\mathbf{\Sigma}\mathbf{W}$$

where $\mathbf{\Sigma} := [\sigma_{ij}]_{(m-1)\times(m-1)}$, $\sigma_{ij} := F(x_q)/\bar{F}(x_q)$, $q = i \wedge j,\ i,j = 1,2,...,m-1$, and $\bar{F}(x) := ax^{-\alpha}\left(1 + bx^{-\beta} + o\left(x^{-\beta}\right)\right)$. By hypothesis $bx^{-\beta} + o\left(x^{-\beta}\right) \simeq 0$, hence $a = x_0$ and $\bar{F}(x_i) \simeq (x_0/x_i)^{\alpha}$, therefore

$$\sigma_{ii} := \frac{F(x_i)}{\bar{F}(x_i)} \simeq \frac{k(i-m)+m}{k(m-i)} = \frac{1}{k}\left(\frac{k(i+m)}{m-i} + \frac{m}{m-i}\right) \simeq \frac{1}{k}\frac{m}{m-i} \text{ (for small values of } k\text{).}$$

Thus, using $\sigma_{ij} := \frac{1}{k}\frac{m}{m-q}$ to construct $\mathbf{\Sigma}$, one can verify that $\mathbf{W}'\mathbf{\Sigma}\mathbf{W}$ yields,

$$Var\left(\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right) \simeq \frac{m}{k}\begin{bmatrix} d & e \\ e & f \end{bmatrix}$$

where

$$\begin{aligned}
d \quad &: = m - 1 + \sum_{j=1}^{m-2}\sum_{i=1}^{j}\frac{1}{m-i} = 2m + o(m), \\
e \quad &: = \sum_{i=1}^{m-1} z_i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{1}{m-i} = -\frac{3m}{\alpha} + o(m) \\
f \quad &: = \frac{z_1}{m-1}\sum_{i=1}^{m-1} z_i + \sum_{j=1}^{m-2} z_{j+1}\sum_{i=1}^{j}\frac{z_i}{m-i} + \sum_{j=1}^{m-1}\frac{z_j}{m-j}\sum_{i=j}^{m-1} z_i = \frac{6m}{\alpha^2} + o(m).
\end{aligned}$$

Note that $d = 2m - 2 - \sum_{i=1}^{m-1} 1/i = 2m + o(m)$. Also in this case there are no closed-form expressions for e and f, but we determine numerically that $e = -3m/\alpha + o(m)$ and $f = 6m/\alpha^2 + o(m)$. ■

Proof of Theorem 1 (i) Let $\tilde{w}_i = \begin{pmatrix} 1 & \tilde{z}_i \end{pmatrix}'$ and $\theta = \begin{pmatrix} \vartheta & \alpha \end{pmatrix}'$. Considering $\tilde{\alpha}^{-1} \xrightarrow{p} \alpha^{-1}$ we have that,

$$\begin{aligned}
\hat{\theta} &= \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i y_i \\
&= \theta + \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\varepsilon_i \\
&= \theta + \left(\frac{1}{m}\sum_{i=1}^{m-1}w_i w_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}w_i\varepsilon_i + o_p(1).
\end{aligned}$$

To justify the third equality consider for example

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m-1}\tilde{z}_i\varepsilon_i &= \frac{1}{m}\sum_{i=1}^{m-1}\frac{1}{\tilde{\alpha}}\log\left(1-u_i\right)\varepsilon_i \\
&= \frac{1}{m}\sum_{i=1}^{m-1}\left(\frac{1}{\alpha}+o_p\left(1\right)\right)\log\left(1-u_i\right)\varepsilon_i \\
&= \frac{1}{m}\sum_{i=1}^{m-1}\frac{1}{\alpha}\log\left(1-u_i\right)\varepsilon_i+\frac{o_p\left(1\right)}{m}\sum_{i=1}^{m-1}\log\left(1-u_i\right)\varepsilon_i \\
&= \frac{1}{m}\sum_{i=1}^{m-1}z_i\varepsilon_i+o_p\left(1\right).
\end{aligned}$$

Now let us focus on the term

$$\frac{1}{m}\sum_{i=1}^{m-1}w_i\varepsilon_i=\frac{1}{m}\sum_{i=1}^{m-1}\begin{pmatrix}\varepsilon_i \\ \alpha^{-1}\log\left(1-u_i\right)\varepsilon_i\end{pmatrix}.$$

By the Glivenko and Canteli theorem (see Wellner, 1977) it follows that $\sup_x\left|\bar{F}_n\left(x\right)-\bar{F}\left(x\right)\right|\xrightarrow{p}0$ as $n\rightarrow\infty$ and, due to the continuity of the log function and the fact that $\lim E\left(\log\left(1-u_i\right)\log\bar{F}_n\left(x_i\right)\right)<\infty$, we have that $\sup_{x_i}\left|\varepsilon\left(x_i\right)\right|\xrightarrow{p}0$ and $\sup_{x_i}\left|\log\left(1-u_i\right)\varepsilon\left(x_i\right)\right|\xrightarrow{p}0$, $\varepsilon\left(x_i\right):=\varepsilon_i=\log\bar{F}_n\left(x_i\right)-\log\bar{F}\left(x_i\right)$ (actually, point wise convergence in probability is enough, in view of the way x_i is generated). It remains to be shown that $m^{-1}\sum_{i=1}^{m-1}w_iw_i'$ converges to a positive definite matrix. This is immediate in view of Lemma 2. Thus, $\hat{\theta}\xrightarrow{p}\theta$ and in particular $\hat{\alpha}\xrightarrow{p}\alpha$.

(ii) Note that,

$$\begin{aligned}
\sqrt{n}\left(\hat{\theta}-\theta\right) &= \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\sqrt{n}\varepsilon_i \\
&= \left(\frac{1}{m}\sum_{i=1}^{m-1}w_iw_i'\right)^{-1}\frac{\sqrt{n}}{m}\sum_{i=1}^{m-1}w_i\varepsilon_i+o_p\left(1\right).
\end{aligned}$$

The second equality needs a brief explanation. The issue is whether the sampling variation of $\tilde{\alpha}$ can (at least asymptotically) be ignored. In this case, the limiting distribution of the OLS estimators is the same as that of the OLS estimators when α is replaced by $\tilde{\alpha}$. Wooldridge (2010, section 6.1.1 and 12.4.2) provides a simple and sufficient condition: $\lim E\left(\left(\partial z_i/\partial\alpha\right)\varepsilon_i\right)=0$. Following previous arguments (see (i) of this proof) it is straightforward to verify this condition.

Since $\sqrt{n}\sum_{i=1}^{m-1}\varepsilon_i \xrightarrow{d}$ Normal (by Lemma 1), $\frac{\sqrt{n}}{m}\sum_{i=1}^{m-1} w_i\varepsilon_i$ will converge in distribution to a proper non-degenerate distribution if $\lim Var\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right)$ converges to a constant positive definite matrix (notice that w_i is a deterministic function). From (ii) of Lemma 2 we have that

$$Var\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right) = \frac{1}{m^2}Var\left(\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right) \to \begin{bmatrix} 1 & -\frac{2}{\alpha} \\ -\frac{2}{\alpha} & \frac{5}{\alpha^2} \end{bmatrix}.$$

Furthermore, from the same Lemma it follows that

$$\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i'\right)^{-1} \longrightarrow \begin{bmatrix} 1 & -\frac{1}{\alpha} \\ -\frac{1}{\alpha} & \frac{2}{\alpha^2} \end{bmatrix}^{-1} = \begin{bmatrix} 2 & \alpha \\ \alpha & \alpha^2 \end{bmatrix}.$$

Therefore

$$\lim Var\left(\sqrt{n}\left(\hat{\theta}-\theta\right)\right) = \begin{bmatrix} 2 & \alpha \\ \alpha & \alpha^2 \end{bmatrix}\begin{bmatrix} 1 & -\frac{2}{\alpha} \\ -\frac{2}{\alpha} & \frac{5}{\alpha^2} \end{bmatrix}\begin{bmatrix} 2 & \alpha \\ \alpha & \alpha^2 \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ \alpha & 2\alpha^2 \end{bmatrix}$$

and, in particular,

$$\lim Var\left(\sqrt{n}\left(\hat{\alpha}-\alpha\right)\right) = 2\alpha^2.$$

■

Proof of Theorem 2

(i) Let $\tilde{w}_i = \begin{pmatrix} 1 & \tilde{z}_i \end{pmatrix}'$, $\theta = \begin{pmatrix} \vartheta & \alpha \end{pmatrix}'$. Note that,

$$\begin{aligned}
\hat{\theta} &= \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i y_i \\
&= \theta + \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\varepsilon_i \\
&\quad + \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\log\left(1+b\left(1-u_i\right)^{\frac{\beta}{\alpha}}x_0^{-\beta}+o\left(x_i^{-\beta}\right)\right) \\
&= \theta + \left(\frac{1}{m}\sum_{i=1}^{m-1}w_i w_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}w_i\varepsilon_i \\
&\quad + \left(\frac{1}{m}\sum_{i=1}^{m-1}w_i w_i'\right)^{-1}\frac{1}{m}\sum_{i=1}^{m-1}w_i\log\left(1+b\left(1-u_i\right)^{\frac{\beta}{\alpha}}x_0^{-\beta}+o\left(x_i^{-\beta}\right)\right) + o_p\left(1\right). \\
&=: \theta + I_1 + I_2 + o_p(1)
\end{aligned}$$

We have seen that I_1 converges in probability to zero (see proof of Proposition 1). Thus, we only focus on the analysis of I_2. Note that,

$$\frac{1}{m}\sum_{i=1}^{m-1} w_i \log\left(1+bx_i^{-\beta}\right) = \frac{1}{m}\sum_{i=1}^{m-1}\begin{pmatrix} \log\left(1+b\left(1-u_i\right)^{\frac{\beta}{\alpha}} x_0^{-\beta} + o\left(x_i^{-\beta}\right)\right) \\ \alpha^{-1}\log\left(1-u_i\right)\log\left(1+b\left(1-u_i\right)^{\frac{\beta}{\alpha}} x_0^{-\beta} + o\left(x_i^{-\beta}\right)\right)\end{pmatrix}. \tag{A.7}$$

Considering $x_0 = F^{-1}(1-k) \simeq c_0 k^{-1/\alpha}$ for small k (where the tails of $\bar{F}$ are well approximated by those of the Pareto Law, $\bar{F}(x) \simeq x^{-\alpha}c_1$). With $k = m/n \to 0$, $\beta > \alpha$, $x_0^{-\beta} = \left(c_0 k^{-1/\alpha}\right)^{-\beta} = c_2 k^{\frac{\beta}{\alpha}} = c_2\left(\frac{m}{n}\right)^{\beta/\alpha}$we have

$$\begin{aligned}
&\left|\frac{1}{m}\sum_{i=1}^{m-1}\log\left(1+b\left(1-u_i\right)^{\frac{\beta}{\alpha}} x_0^{-\beta} + o\left(x_i^{-\beta}\right)\right)\right| \quad \leq \\
&\leq \quad \left|\frac{1}{m}\sum_{i=1}^{m-1}\log\left(1+bx_0^{-\beta} + o\left(x_i^{-\beta}\right)\right)\right| \\
&\simeq \quad \left|\frac{1}{m}\sum_{i=1}^{m-1} bx_0^{-\beta} + o\left(x_i^{-<\beta}\right)\right| \text{ (since } x_0^{-\beta} \to 0) \\
&\leq \quad \left|\frac{1}{m}\sum_{i=1}^{m-1}(b+1)\,x_0^{-\beta}\right| \text{ (since } o\left(x_i^{-\beta}\right) < x_0^{-\beta}) \\
&= \quad \left|\frac{m-1}{m}(b+1)\,c_2\left(\frac{m}{n}\right)^{\beta/\alpha}\right| \\
&\leq \quad (b+1)\,c_2 m^{\frac{\beta}{\alpha}(1-\gamma)} \text{ (since } n = m^{\gamma})
\end{aligned}$$

and this expression converges to zero as $m/n \to 0$ with $\gamma > 1$ (or as $m, n \to \infty$ and $n = m^{\gamma}$). Using the same reasoning, the second element of the vector (A.7) converges to zero as well.

(ii) We have that,

$$\begin{aligned}
\sqrt{m}\left(\hat{\theta}-\theta\right) &= \left(\frac{1}{m}\sum_{i=1}^{m-1}\tilde{w}_i\tilde{w}_i'\right)^{-1}\frac{\sqrt{m}}{m}\sum_{i=1}^{m-1}\tilde{w}_i\varepsilon_i \\
&= \left(\frac{1}{m}\sum_{i=1}^{m-1}w_i w_i'\right)^{-1}\frac{1}{\sqrt{nm}}\sum_{i=1}^{m-1}w_i\sqrt{n}\varepsilon_i + o_p(1).
\end{aligned}$$

Since $\sqrt{n}\sum_{i=1}^{m-1}\varepsilon_i \xrightarrow{d}$ Normal (by Lemma 1), $\frac{1}{\sqrt{nm}}\sum_{i=1}^{m-1} w_i\varepsilon_i$ converges in distribution to a proper non-degenerate distribution if $\lim Var\left(\frac{1}{\sqrt{nm}}\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right)$

converges to a constant positive definite matrix (notice that w_i is a deterministic function). By Lemma 2 (ii) we have

$$\begin{aligned} Var\left(\frac{1}{\sqrt{nm}}\sum_{i=1}^{m-1} w_i\varepsilon_i\right) &= \frac{1}{nm}Var\left(\sum_{i=1}^{m-1} w_i\sqrt{n}\varepsilon_i\right) \\ &= \frac{1}{nm}\frac{m}{k}\left[\begin{array}{cc} 2m+o\left(m\right) & \frac{-3m}{\alpha}+o\left(m\right) \\ \frac{-3m^2}{\alpha}+o\left(m\right) & \frac{6m}{\alpha^2}+o\left(m\right) \end{array}\right] \rightarrow \left[\begin{array}{cc} 2 & \frac{-3}{\alpha} \\ \frac{-3}{\alpha} & \frac{6}{\alpha^2} \end{array}\right] \end{aligned}$$

with $k = m/n$. On the other hand, by the same lemma,

$$\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i'\right)^{-1} \longrightarrow \left[\begin{array}{cc} 2 & \alpha \\ \alpha & \alpha^2 \end{array}\right].$$

Therefore

$$\lim Var\left(\sqrt{m}\left(\hat{\theta}-\theta\right)\right) = \left[\begin{array}{cc} 2 & \alpha \\ \alpha & \alpha^2 \end{array}\right]\left[\begin{array}{cc} 2 & \frac{-3}{\alpha} \\ \frac{-3}{\alpha} & \frac{6}{\alpha^2} \end{array}\right]\left[\begin{array}{cc} 2 & \alpha \\ \alpha & \alpha^2 \end{array}\right] = \left[\begin{array}{cc} 2 & \alpha \\ \alpha & 2\alpha^2 \end{array}\right]$$

and

$$\lim Var\left(\sqrt{m}\left(\hat{\alpha}-\alpha\right)\right) = 2\alpha^2.$$

■

Proof of Theorem 3

Consider $\theta := \left(\begin{array}{cc} \vartheta & \alpha \end{array}\right)'$. **(i)** We have established in the proof of Theorem 2 that $\hat{\theta} = \theta + I_1 + I_2 + o_p(1)$. We proved that $I_2 \to 0$ (in deterministic sense). Thus, we focus on $I_1 = \left(\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i'\right)^{-1} \frac{1}{m}\sum_{i=1}^{m-1} w_i\varepsilon_i$. Given that a stationary process that satisfies the strong mixing condition $\alpha\left(n\right) \to 0$, is also a stationary ergodic process (Rosenblatt, 1978) we may apply the Glivenko and Canteli theorem for stationary ergodic sequences, which guarantees that $\sup_x \left|\bar{F}_n\left(x\right) - \bar{F}\left(x\right)\right| \xrightarrow{p} 0$ as $n \to \infty$. The rest of the proof is similar to that of proof of Theorem 1, and thus $I_1 \xrightarrow{p} 0$, which implies $\hat{\theta} \xrightarrow{p} \theta$. **(ii)** Under the conditions of the theorem, Deo (1973) has proved that

$$\sqrt{n}\left(\begin{array}{c} \bar{F}_n\left(x_1\right) - \bar{F}\left(x_1\right) \\ \bar{F}_n\left(x_2\right) - \bar{F}\left(x_2\right) \\ \vdots \\ \bar{F}_n\left(x_{m-1}\right) - \bar{F}\left(x_{m-1}\right) \end{array}\right) \xrightarrow{d} N\left(0, \mathbf{A}\right)$$

where $\mathbf{A}_{ij} = \lim E\left(g_1\left(x_i\right)g_1\left(x_j\right)\right) + \sum_{n\geq 2}\lim E\left(g_1\left(x_i\right)g_n\left(x_j\right)\right) + \sum_{n\geq 2}\lim E\left(g_n\left(x_i\right)g_1\left(x_j\right)\right)$, $g_n\left(x\right) = \mathcal{I}_{\{X_n\leq x\}} - F\left(x\right)$ and these two series converges absolutely for $x_i,\ x_j \in \mathbb{R}$.

Using the same arguments as in the proof of Lemma 1 one has $\sqrt{n}\varepsilon \xrightarrow{d} N\left(0,\mathbf{\Sigma}\right)$ where $\varepsilon := \left(\varepsilon_1, ..., \varepsilon_{m-1}\right)'$, $\varepsilon_i := \log \bar{F}_n\left(x_i\right) - \log \bar{F}\left(x_i\right)$ and $\mathbf{\Sigma} := \left[\sigma_{ij}\right]_{(m-1)\times(m-1)}$. The rest of the proof is similar to that of Theorem 2, but now $J_m = \lim Var\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i\varepsilon_i\right)$ has an unknown expression given the dependence structure of the sequence $\{\varepsilon_i\}$. In the case of i.i.d. sequence $\{\varepsilon_i\}$ we have seen that $J_m^{iid} = \lim Var\left(\frac{1}{m}\sum_{i=1}^{m-1} w_i\varepsilon_i\right)$ is a matrix with elements $O\left(1\right)$. Changing from the i.i.d. hypothesis to the dependence case does not alter the order of magnitude of J_n, since in both cases $\sqrt{n}\varepsilon_i$ has a limit distribution that is independent of m. Obviously, we expected $J_n - J_n^{iid}$ to be a positive-semidefinite matrix. As a result we have

$$\sqrt{m}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(\mathbf{0}, Avar\left(\hat{\theta}\right)\right)$$

$$\begin{aligned} Avar\left(\hat{\theta}\right) &= \lim_{m\to\infty} \sqrt{m}\left(\hat{\theta} - \theta\right) = \lim_{m\to\infty} \left(\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i'\right)^{-1} J_m \left(\frac{1}{m}\sum_{i=1}^{m-1} w_i w_i'\right)^{-1} \\ &= \left[\begin{array}{cc} 2 & \alpha \\ \alpha & \alpha^2 \end{array}\right] \lim_{m\to\infty} J_m \left[\begin{array}{cc} 2 & \alpha \\ \alpha & \alpha^2 \end{array}\right]. \blacksquare \end{aligned}$$

Proof of Theorem 4

Under the hypotheses defined, the conditions of the Theorem 1 of Andrews and Monahan (1992) (AM) hold. In fact, Assumption A of AM is implied by the α-mixing and the moment condition (we note that $\left\|V_t V_{t-j}'\right\|$, with $V_t = w_i\varepsilon_i$, has moments of any order given that V_t is formed by a sequence of a deterministic sequence $\left(1, z_i\right)'$ and a random variable with asymptotic normal distribution). The same reasoning applies to $\lim E\left(\left\|\partial V_t/\partial\theta'\right\|^2\right)$; (see Assumption B of AM). Also $\sqrt{m}\left(\hat{\theta} - \theta\right) = O\left(1\right)$, as seen in theorem 3; see also Assumption B of AM. The parameter $\hat{S}_m$ and the kernel as defined above satisfy respectively the Assumption C of AM and the condition defined in Theorem 1 of AM. $\blacksquare$

Table 1: Descriptive Statistics of the Exchange Rate Returns

		ARS	AUD	BRL	CAD	CHF	CLP	COP	DKK	EUR	GBP	HKD	KRW	MXP	NOK	NZD	PLN	SEK	SGD	UAH	YEN	ZAR
I	μ	0.14	-0.01	0.07	-0.00	-0.03	0.02	0.07	-0.00	-0.00	-0.00	0.00	0.02	0.10	0.00	-0.02	0.03	0.00	-0.01	0.17	0.00	0.08
	Std	0.15	0.15	0.18	0.10	0.13	0.11	0.12	0.12	0.11	0.10	0.01	0.17	0.17	0.14	0.15	0.15	0.14	0.07	0.21	0.13	0.18
	Skw	18.1	0.75	0.47	-0.10	-0.79	0.58	0.38	-0.19	-0.18	0.04	-2.72	-0.76	2.82	-0.03	0.37	0.19	-0.18	-0.40	10.95	-0.46	0.32
	Kur	700.2	16.3	23.3	10.1	23.3	10.7	13.0	5.56	5.61	7.45	68.0	108.6	97.0	8.32	9.14	8.76	6.64	14.2	416.9	8.13	10.2
II	μ	0.00	0.01	0.15	0.02	0.02	0.05	0.20	0.02	0.03	-0.02	0.00	0.08	0.26	0.02	0.02	0.15	0.01	0.01	0.43	-0.02	0.14
	Std	0.01	0.11	0.15	0.06	0.13	0.07	0.09	0.11	0.101	0.09	0.00	0.24	0.26	0.11	0.11	0.10	0.11	0.08	0.32	0.15	0.12
	Skw	0.18	-0.32	2.30	-0.15	-0.45	-0.06	0.69	-0.36	-0.33	-0.06	-1.56	-0.68	2.65	-0.24	-0.40	0.67	-0.15	-0.96	0.00	-0.72	0.24
	Kur	25.4	9.82	60.3	6.13	5.80	13.8	18.1	4.88	4.91	5.30	39.0	81.1	64.0	10.4	9.22	10.7	5.14	17.5	8.78	8.93	19.3
III	μ	0.20	-0.05	-0.00	-0.07	-0.06	-0.01	0.01	-0.07	-0.07	-0.04	0.00	-0.03	0.03	-0.07	-0.07	-0.09	-0.05	-0.03	-0.01	0.02	0.02
	Std	0.24	0.13	0.19	0.09	0.12	0.11	0.10	0.11	0.11	0.10	0.01	0.08	0.09	0.12	0.14	0.12	0.12	0.05	0.07	0.11	0.19
	Skw	11.8	0.55	-0.19	0.08	-0.13	0.19	0.22	-0.12	-0.12	-0.02	-5.14	0.31	0.33	0.07	0.64	0.36	-0.04	-0.14	1.59	-0.35	0.10
	Kur	294.9	6.12	20.3	4.21	3.83	5.37	11.8	4.09	4.09	3.77	117.3	5.24	5.65	3.78	5.97	5.33	3.70	6.31	39.9	4.58	8.03
IV	μ	0.20	0.02	0.09	0.05	-0.04	0.04	0.03	0.05	0.05	0.05	-0.00	0.03	0.06	0.07	0.01	0.08	0.05	-0.01	0.32	0.01	0.11
	Std	0.06	0.19	0.19	0.13	0.15	0.14	0.15	0.13	0.13	0.12	0.01	0.17	0.15	0.17	0.18	0.21	0.17	0.07	0.29	0.13	0.21
	Skw	21.4	0.91	0.44	-0.17	-1.30	0.76	0.37	-0.18	-0.17	0.07	-0.38	-0.72	0.67	-0.04	0.32	0.06	-0.25	0.20	9.98	-0.18	0.51
	Kur	700.6	15.8	12.8	8.67	37.4	9.64	9.81	6.56	6.51	8.20	15.5	45.1	13.9	7.63	8.70	6.42	6.27	7.40	281.5	7.71	8.93

Note: I = 1994 - 2015 (T=5510); II = 1994 - 1999 (T=1565); III = 2000 - 20007 (T=2086) and IV = 2008 - 2015 (T = 1859) and where μ is the annualized mean return; std is the annualized standard deviation; and skw and kur refer to the skweness and kurtosis of the series.

Table 2: Right Tail Index of the Exchange Rate Returns Series

	1994 - 2015		1994 - 1999		2000 - 2007		2008 - 2015	
	$\widehat{\alpha}_H$	$\widehat{\alpha}_{Pareto}$	$\widehat{\alpha}_H$	$\widehat{\alpha}_{Pareto}$	$\widehat{\alpha}_H$	$\widehat{\alpha}_{Pareto}$	$\widehat{\alpha}_H$	$\widehat{\alpha}_{Pareto}$
ARS $_{\mathcal{FM}}$	1.070	1.098	1.689	1.564	1.054	1.070	1.781	1.806
	(0.041)	(0.045)	(0.070)	(0.100)	(0.064)	(0.072)	(0.095)	(0.098)
		[0.032]		[1.115]		[0.094]		[0.036]
AUD $_{\mathcal{DM}}$	2.249	2.220	2.358	2.370	2.521	2.467	2.124	2.029
	(0.055)	(0.063)	(0.109)	(0.123)	(0.092)	(0.109)	(0.094)	(0.104)
		[1.502]		[0.122]		[0.315]		[0.060]
BRL $_{\mathcal{EM}}$	1.872	1.863	1.191	1.061	2.150	2.140	2.044	2.008
	(0.050)	(0.059)	(0.074)	(0.086)	(0.086)	(0.101)	(0.091)	(0.104)
		[0.060]		[0.072]		[0.295]		[0.561]
CAD $_{\mathcal{DM}}$	2.216	2.309	2.633	2.419	2.680	2.734	2.396	2.304
	(0.054)	(0.065)	(0.111)	(0.124)	(0.096)	(0.114)	(0.096)	(0.111)
		[0.253]		[0.246]		[1.440]		[0.593]
CHF $_{\mathcal{DM}}$	2.788	2.783	2.747	2.694	3.124	2.863	2.561	2.461
	(0.058)	(0.071)	(0.113)	(0.131)	(0.096)	(0.117)	(0.098)	(0.115)
		[0.169]		[0.225]		[0.215]		[0.168]
CLP $_{\mathcal{EM}}$	2.075	2.153	2.164	2.206	2.462	2.570	2.094	2.014
	(0.053)	(0.063)	(0.098)	(0.119)	(0.093)	(0.110)	(0.092)	(0.104)
		[0.160]		[0.401]		[3.030]		[0.146]
COP $_{\mathcal{EM}}$	2.012	1.884	1.936	1.971	1.943	1.950	1.988	1.880
	(0.050)	(0.058)	(0.099)	(0.112)	(0.081)	(0.097)	(0.089)	(0.101)
		[0.112]		[0.148]		[0.198]		[0.078]
DKK $_{\mathcal{DM}}$	2.704	2.926	2.863	2.948	2.725	2.879	2.687	2.653
	(0.057)	(0.073)	(0.114)	(0.137)	(0.093)	(0.117)	(0.094)	(0.120)
		[1.253]		[2.531]		[3.204]		[1.269]
EUR $_{\mathcal{DM}}$	2.646	2.899	2.884	2.850	2.696	2.846	2.588	2.654
	(0.056)	(0.073)	(0.111)	(0.135)	(0.093)	(0.117)	(0.093)	(0.119)
		[1.237]		[1.651]		[1.750]		[1.425]
GBP $_{\mathcal{DM}}$	2.791	2.574	2.748	2.542	3.161	2.949	2.608	2.357
	(0.056)	(0.068)	(0.106)	(0.127)	(0.095)	(0.119)	(0.098)	(0.113)
		[0.265]		[0.179]		[0.406]		[0.160]
HKD $_{\mathcal{DM}}$	1.534	1.564	1.600	1.716	1.504	1.352	1.657	1.658
	(0.045)	(0.053)	(0.098)	(0.104)	(0.066)	(0.081)	(0.080)	(0.094)
		[0.050]		[0.143]		[0.091]		[1.036]
KRW $_{\mathcal{EM}}$	1.567	1.512	1.051	1.008	2.285	1.938	1.803	1.813
	(0.046)	(0.052)	(0.075)	(0.080)	(0.081)	(0.096)	(0.086)	(0.099)
		[0.067]		[0.036]		[0.084]		[0.240]
MXP $_{\mathcal{EM}}$	1.813	1.731	1.269	1.276	2.435	2.538	2.183	2.103
	(0.052)	(0.056)	(0.086)	(0.090)	(0.095)	(0.110)	(0.092)	(0.106)
		[0.125]		[0.029]		[0.441]		[0.106]
NOK $_{\mathcal{DM}}$	2.470	2.450	2.686	2.644	3.017	2.774	2.177	2.296
	(0.057)	(0.067)	(0.106)	(0.130)	(0.096)	(0.115)	(0.096)	(0.111)
		[0.136]		[0.200]		[0.193]		[0.106]
NZD $_{\mathcal{DM}}$	2.257	2.213	2.379	2.214	2.280	2.336	2.253	2.180
	(0.054)	(0.063)	(0.102)	(0.119)	(0.088)	(0.106)	(0.094)	(0.108)
		[0.100]		[0.156]		[0.319]		[0.085]

Table 2 (Cont.): Right Tail Index of the Exchange Rate Returns Series

PLN $\mathcal{DM}$	2.028	2.113	2.074	2.025	2.549	2.395	2.463	2.253
	(0.052)	(0.063)	(0.111)	(0.125)	(0.087)	(0.107)	(0.092)	(0.110)
		[1.231]		[0.321]		[0.371]		[1.101]
SEK $\mathcal{DM}$	2.614	2.456	2.656	2.604	2.980	3.068	2.626	2.284
	(0.057)	(0.067)	(0.111)	(0.129)	(0.097)	(0.121)	(0.095)	(0.110)
		[0.160]		[0.254]		[1.425]		[0.108]
SGD $\mathcal{DM}$	2.023	2.028	1.743	1.623	2.628	2.569	2.165	2.096
	(0.052)	(0.061)	(0.087)	(0.102)	(0.093)	(0.111)	(0.092)	(0.106)
		[0.065]		[0.124]		[0.243]		[0.088]
UAH $\mathcal{FM}$	0.932	1.085	1.641	1.583	1.294	1.245	0.932	1.036
	(0.042)	(0.051)	(0.227)	(0.237)	(0.066)	(0.077)	(0.062)	(0.075)
		[0.422]		[0.512]		[1.662]		[0.206]
YEN $\mathcal{DM}$	2.673	2.561	2.623	2.387	3.390	2.876	2.432	2.343
	(0.058)	(0.068)	(0.106)	(0.124)	(0.098)	(0.117)	(0.098)	(0.112)
		[0.083]		[0.081]		[0.269]		[0.638]
ZAR $\mathcal{EM}$	2.103	2.267	1.596	1.539	2.195	2.370	2.346	2.452
	(0.051)	(0.064)	(0.087)	(0.099)	(0.086)	(0.107)	(0.099)	(0.115)
		[2.753]		[0.057]		[8.597]		[0.095]

Note: DM, EM and FM correspond to developed, emerging and frontier markets, respectively. Numbers in parenthesis are the iid standard errors, and in square brackets the robust standard errors.

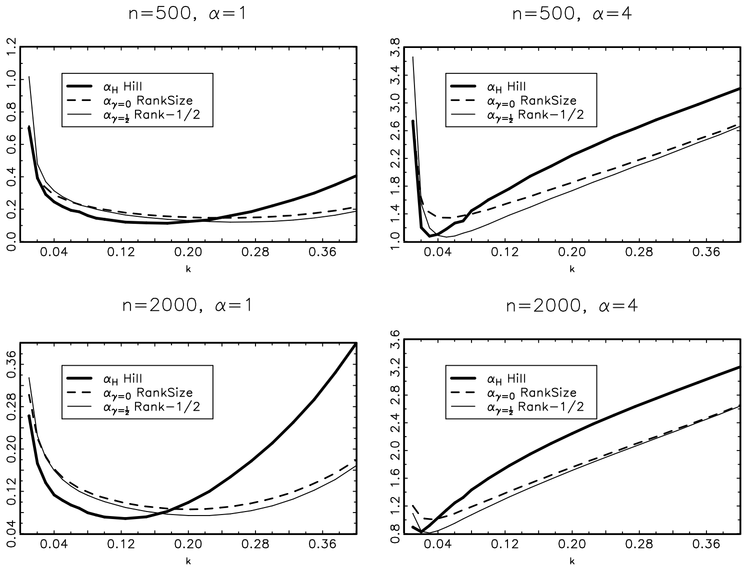

Figure 1: DGP: Student-t Distribution

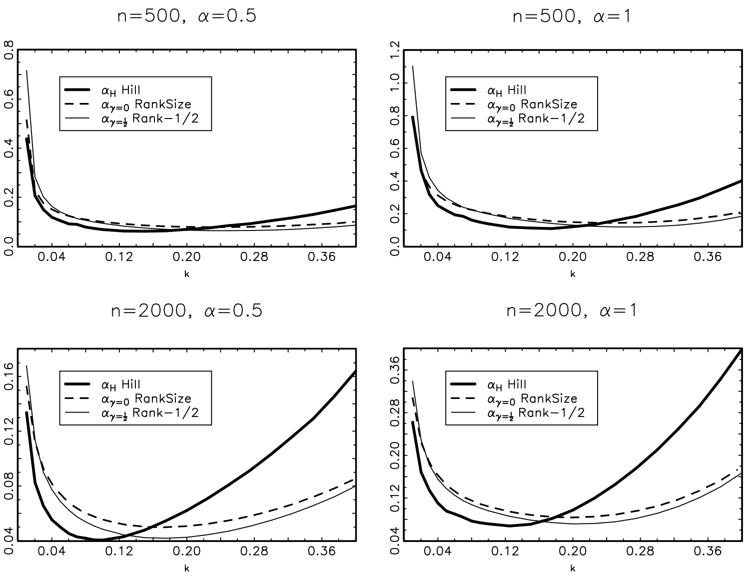

Figure 2: DGP: alpha-Stable Distribution

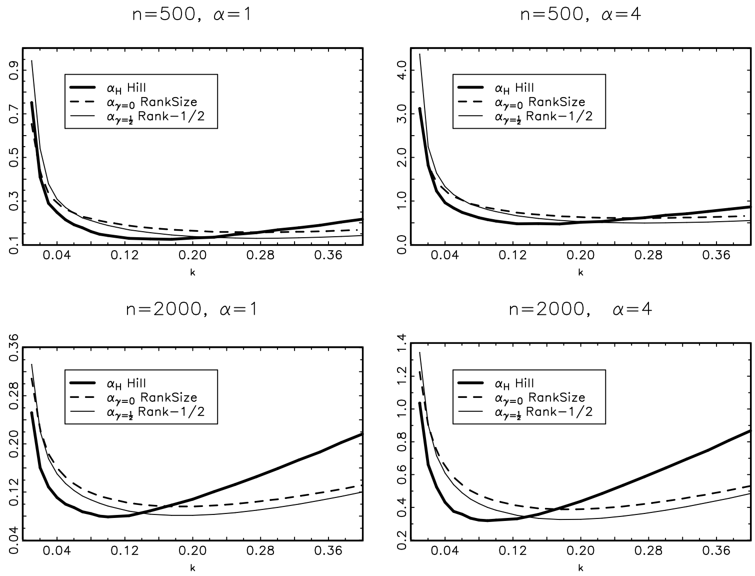

Figure 3: DGP: Burr Distribution

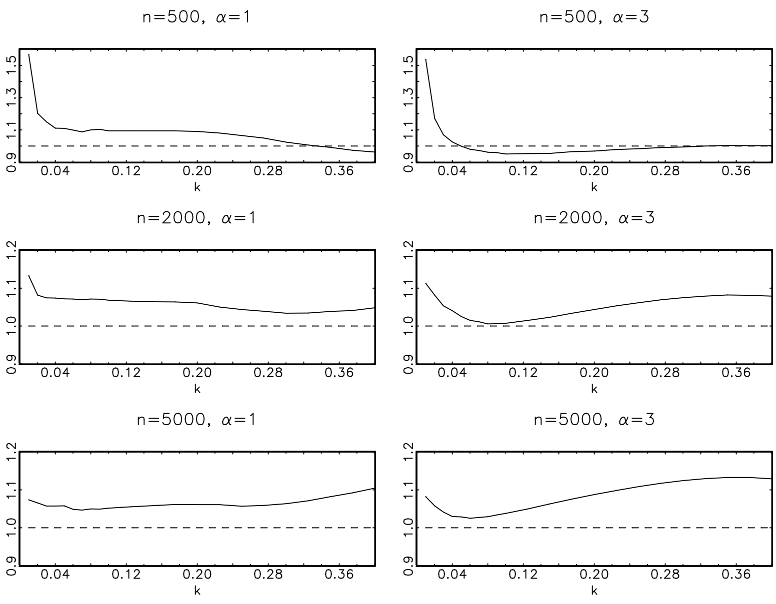

Figure 4: Ratio $RMSE\left(\widehat{\alpha}_{\gamma=1/2}\right)/RMSE\left(\widehat{\alpha}_{Pareto}\right)$. DGT: Student-t Distribution

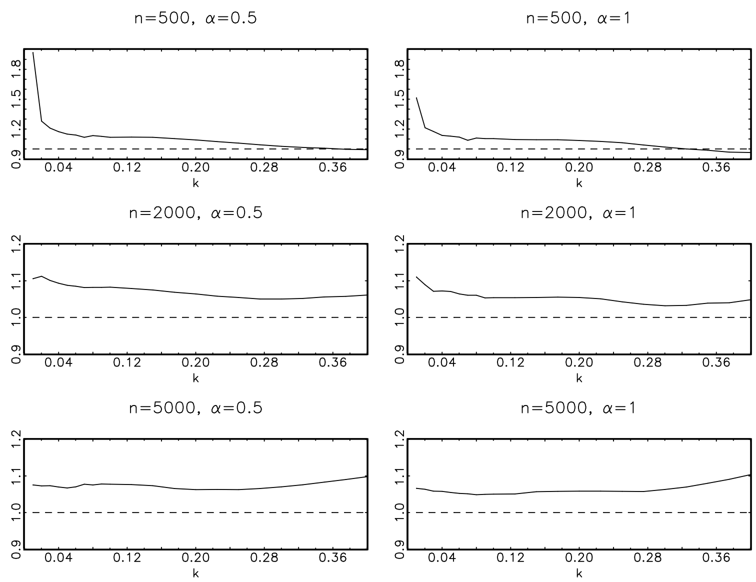

Figure 5: Ratio $RMSE\left(\widehat{\alpha}_{\gamma=1/2}\right)/RMSE\left(\widehat{\alpha}_{Pareto}\right)$. DGT: alpha-Stable Distribution

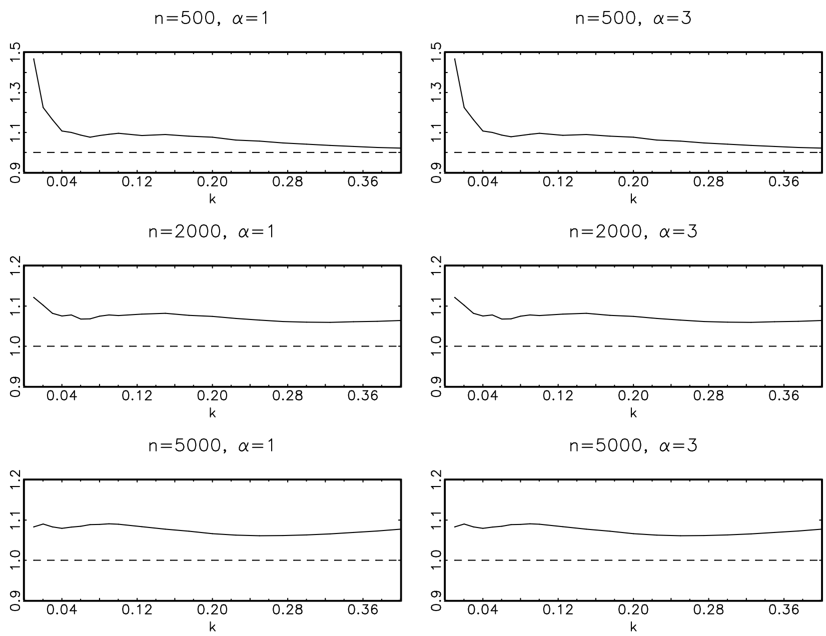

Figure 6: Ratio $RMSE\left(\widehat{\alpha}_{\gamma=1/2}\right)/RMSE\left(\widehat{\alpha}_{Pareto}\right)$. DGT: Burr Distribution

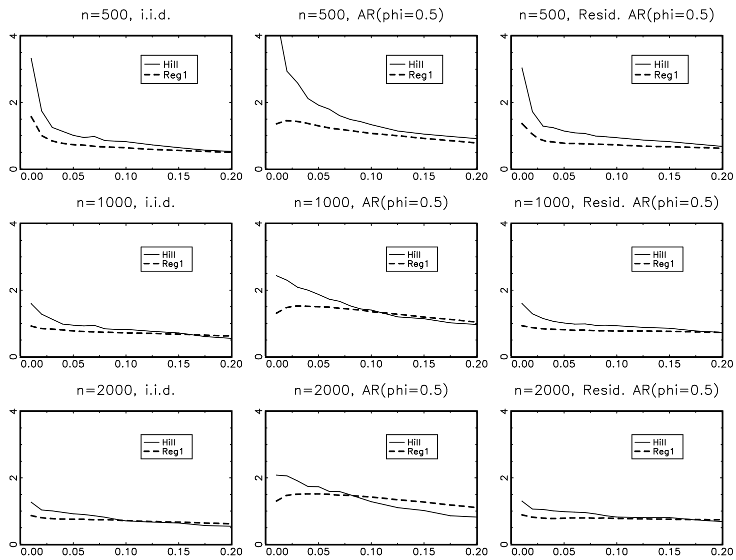

Figure 7: Ratios between the true and the asymptotic variance under the i.i.d. hypothesis. Two estimators are compared: $\hat{\alpha}_{Hill}$ (solid line) and $\hat{\alpha}_{Pareto}$ (dashed line). DGP: Burr distribution with $\alpha = 3$.

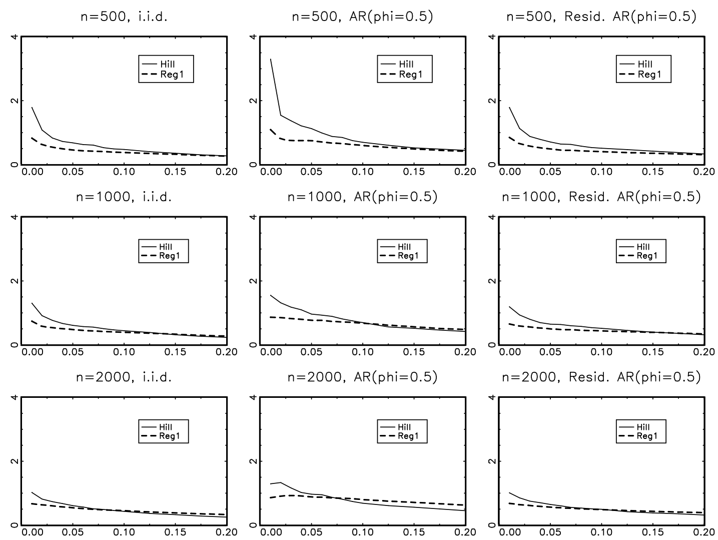

Figure 8: Ratios between the true and the asymptotic variance under i.i.d. hypothesis. Two estimators are compared: $\hat{\alpha}_{Hill}$ (solid line) and $\hat{\alpha}_{Pareto}$ (dashed line). DGP: Student-t distribution with $\alpha = 3$.

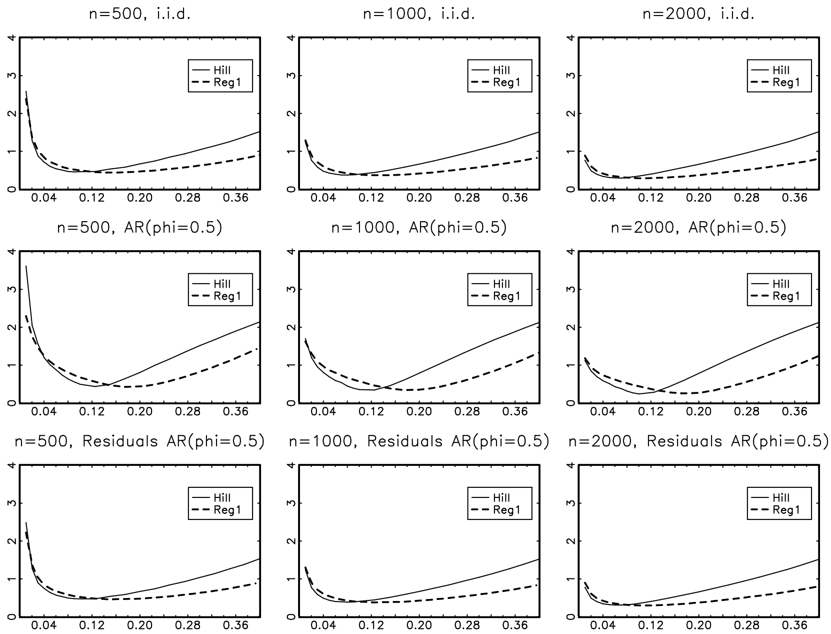

Figure 9: RMSE of $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$. DGP: Burr distribution with $\alpha = 3$.

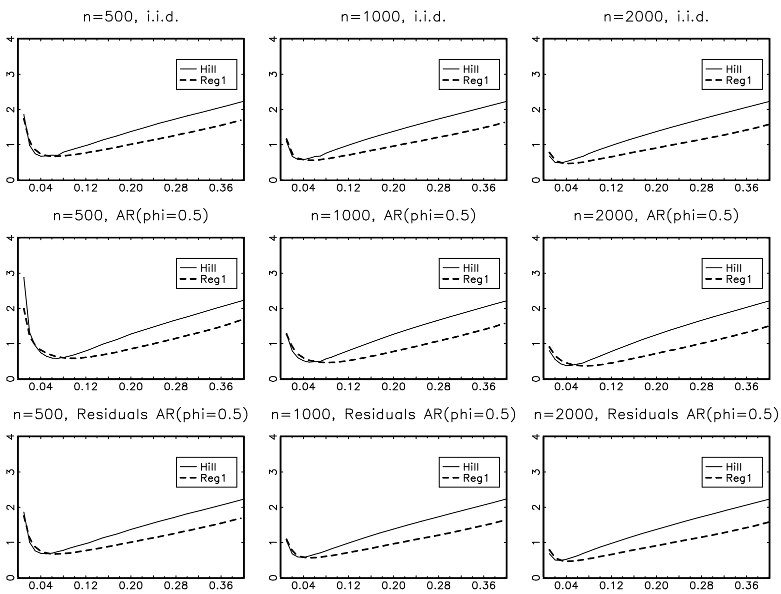

Figure 10: RMSE of $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$. DGP: Student-t distribution with $\alpha = 3$.

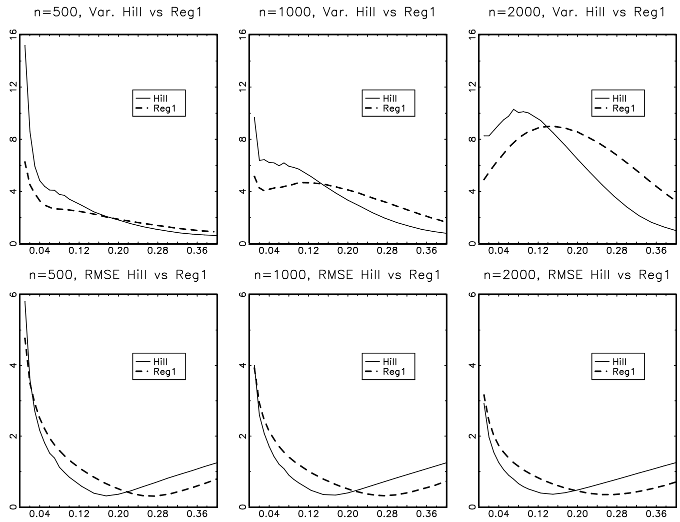

Figure 11: Behaviour of $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$ in the presence of an IGARCH process with Gaussian innovations. The marginal distribution has a tail index of $\alpha = 2$. First Row: Ratio between the true and the theoretical variance of $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$. Second Row: RMSE of $\hat{\alpha}_{Hill}$ and $\hat{\alpha}_{Pareto}$.

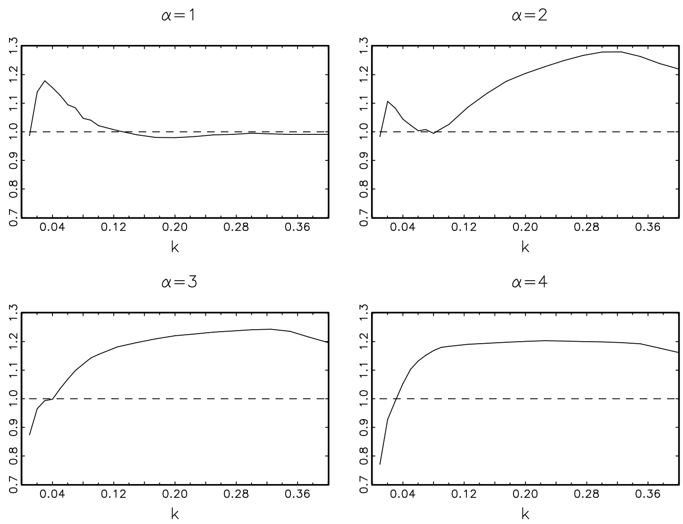

Figure 12: Ratio $RMSE\left(\widehat{\alpha}_{Pareto}\right)/RMSE\left(\widehat{\alpha}_{Student}\right)$. DGT: Student-t Distribution

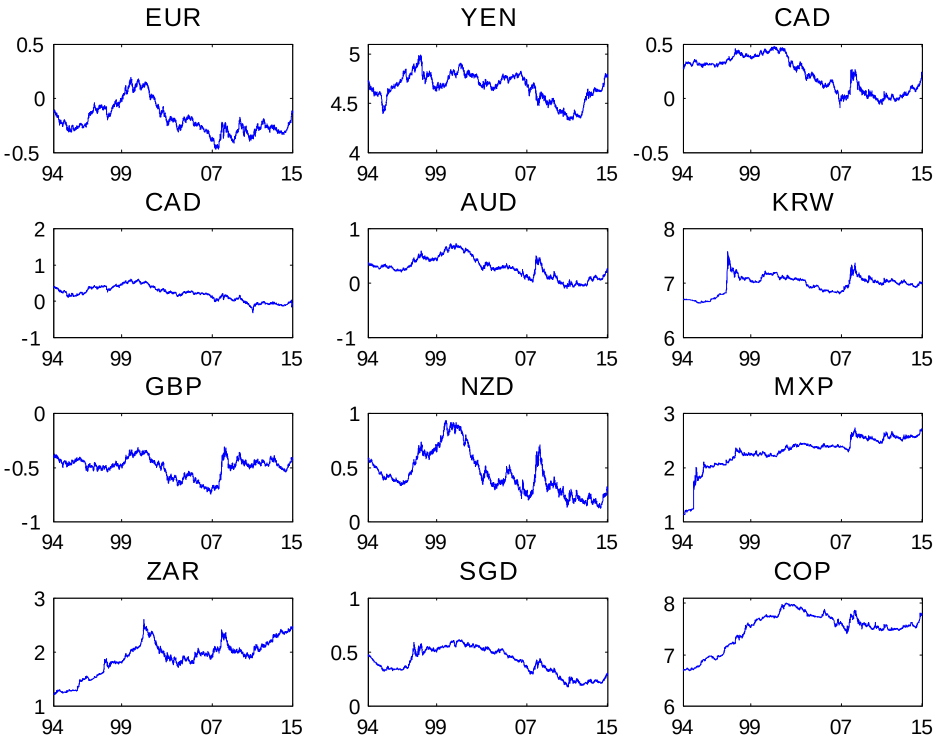

Figure 13: Levels of currency to USD

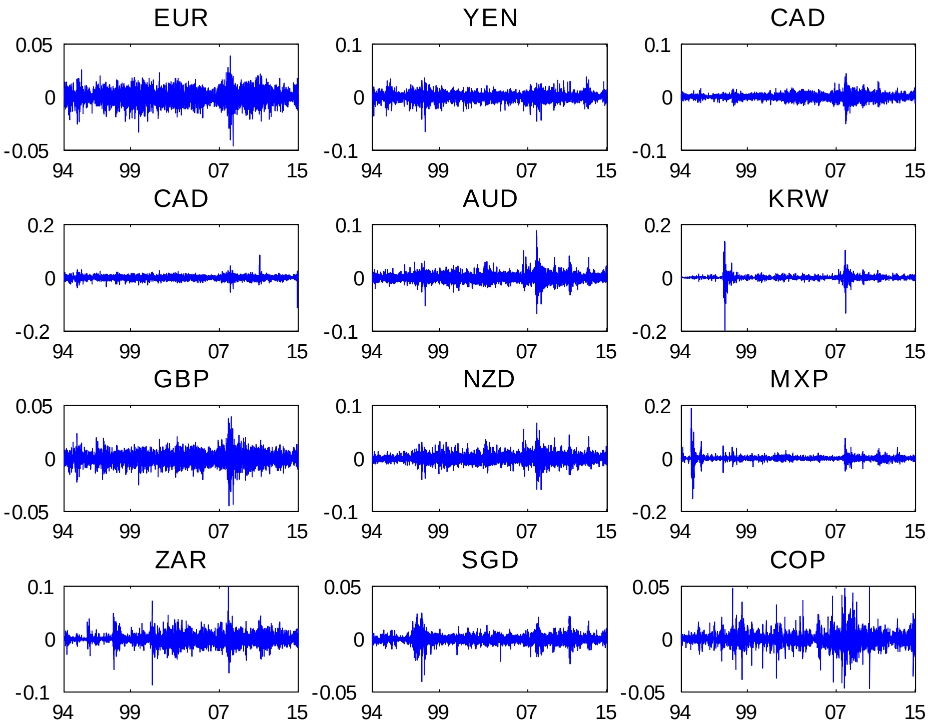

Figure 14: Exchange rate returns

WORKING PAPERS

2013

2014

1|14 Autoregressive augmentation of MIDAS regressions
Cláudia Duarte

2|14 The risk-taking channel of monetary policy – exploring all avenues
Diana Bonfim | Carla Soares

3|14 Global value chains: Surveying drivers, measures and impacts
João Amador | Sónia Cabral

4|14 Has US household deleveraging ended? a model-based estimate of equilibrium debt
Bruno Albuquerque | Ursel Baumann | Georgi Krustev

5|14 The weather effect: estimating the effect of voter turnout on electoral outcomes in italy
Alessandro Sforza

6|14 Persistence in the banking industry: fractional integration and breaks in memory
Uwe Hassler | Paulo M.M. Rodrigues | Antonio Rubia

7|14 Financial integration and the great leveraging
Daniel Carvalho

8|14 Euro area structural reforms in times of a global crisis
Sandra Gomes

9|14 Labour demand research: towards a better match between better theory and better data
John T. Addison | Pedro Portugal | José Varejão

10|14 Capital inflows and euro area long-term interest rates
Daniel Carvalho | Michael Fidora

11|14 Misallocation and productivity in the lead up to the Eurozone crisis
Daniel A. Dias | Carlos Robalo Marquesz | Christine Richmond

12|14 Global value chains: a view from the euro area
João Amador | Rita Cappariello | Robert Stehrer

13|14 A dynamic quantitative macroeconomic model of bank runs
Elena Mattana | Ettore Panetti

14|14 Fiscal devaluation in the euro area: a model-based analysis
S. Gomes | P. Jacquinot | M. Pisani

15|14 Exports and domestic demand pressure: a dynamic panel data model for the euro area countries
Elena Bobeica | Paulo Soares Esteves | António Rua | Karsten Staehr

16|14 Real-time nowcasting the US output gap: singular spectrum analysis at work
Miguel de Carvalho | António Rua

2015

1|15 Unpleasant debt dynamics: can fiscal consolidations raise debt ratios?
Gabriela Castro | Ricardo M. Félix | Paulo Júlio | José R. Maria

2|15 Macroeconomic forecasting starting from survey nowcasts
João Valle e Azevedo | Inês Gonçalves

3|15 Capital regulation in a macroeconomic model with three layers of default
Laurent Clerc | Alexis Derviz | Caterina Mendicino | Stephane Moyen | Kalin Nikolov | Livio Stracca | Javier Suarez | Alexandros P. Vardoulakis

4|15 Expectation-driven cycles: time-varying effects
Antonello D'Agostino | Caterina Mendicino

5|15 Seriously strengthening the tax-benefit link
Pedro Portugal | Pedro S. Raposo

6|15 Unions and collective bargaining in the wake of the great recession
John T. Addison | Pedro Portugal | Hugo Vilares

7|15 Covariate-augmented unit root tests with mixed-frequency data
Cláudia Duarte

8|15 Financial fragmentation shocks
Gabriela Castro | José R. Maria | Paulo Júlio | Ricardo M. Félix

9|15 Central bank interventions, demand for collateral, and sovereign borrowing cost
Luís Fonseca | Matteo Crosignani | Miguel Faria-e-Castro

10|15 Income smoothing mechanisms after labor market transitions
Nuno Alves | Carlos Martins

11|15 Decomposing the wage losses of displaced workers: the role of the reallocation of workers into firms and job titles
Anabela Carneiro | Pedro Raposo | Pedro Portugal

12|15 Sources of the union wage gap: results from high-dimensional fixed effects regression models
John T. Addison | Pedro Portugal | Hugo Vilares

13|15 Assessing european firms' exports and productivity distributions: the compnet trade module
Antoine Berthou | Emmanuel Dhyne | Matteo Bugamelli | Ana-Maria Cazacu | Calin-Vlad Demian | Peter Harasztosi | Tibor Lalinsky | Jaanika Meriküll | Filippo Oropallo | Ana Cristina Soares

14|15 A new regression-based tail index estimator: an application to exchange rates
João Nicolau | Paulo M. M. Rodrigues