

Article

A New Family of Consistent and Asymptotically-Normal Estimators for the Extremal Index

Jose Olmo

Department of Economics, University of Southampton, Bld 58 (Murray Bld), Highfield Campus, Southampton SO17 1BJ, UK; E-Mail: J.B.Olmo@soton.ac.uk; Tel.: +44-238-059-5000

Academic Editor: Gabriel Montes-Rojas

Received: 14 July 2015 / Accepted: 7 August 2015 / Published: 28 August 2015

Abstract: The extremal index (θ) is the key parameter for extending extreme value theory results from i.i.d. to stationary sequences. One important property of this parameter is that its inverse determines the degree of clustering in the extremes. This article introduces a novel interpretation of the extremal index as a limiting probability characterized by two Poisson processes and a simple family of estimators derived from this new characterization. Unlike most estimators for θ in the literature, this estimator is consistent, asymptotically normal and very stable across partitions of the sample. Further, we show in an extensive simulation study that this estimator outperforms in finite samples the logs, blocks and runs estimation methods. Finally, we apply this new estimator to test for clustering of extremes in monthly time series of unemployment growth and inflation rates and conclude that runs of large unemployment rates are more prolonged than periods of high inflation.

Keywords: asymptotic theory; clustering of extremes; extremal index; extreme value theory; order statistics

JEL classifications: C14, C18, C4

1. Introduction

There are different interpretations of the extremal index (θ) in the literature on extreme value theory for weakly dependent processes. This concept, originated in papers by Loynes [1] and O'Brien [2] and developed in detail by Leadbetter [3], reflects the effect of clustering of extreme observations on the limiting distribution of the maximum. Thus, the first author showed that

$$P\{M_{1,n} \leq u_n\} \approx F^{n\theta}(u_n), \quad (1)$$

with $0 \leq \theta \leq 1$, $M_{1,n} = \max\{X_1, \dots, X_n\}$, u_n a normalizing sequence and F the unconditional distribution function of a sequence of weakly dependent random variables $\{X_i, i \geq 1\}$; see also [4,5]. In a similar spirit, [6] showed that the presence of clustering affected the limiting distribution of block maxima, *i.e.*,

$$P\{M_{2,r_n} \leq u_n | X_1 > u_n\} \longrightarrow \theta, \quad (2)$$

with $M_{2,r_n} = \max\{X_2, \dots, X_{r_n}\}$, r_n determining a partition of the sample of length n , such that $r_n \rightarrow \infty$ and $r_n = o(n)$.

Alternatively, Leadbetter [3] showed that for stationary sequences exhibiting short range dependence, the inverse of the extremal index is the limiting mean number of exceedances of u_n in an interval of length r_n . This result mathematically reads as follows

$$E \left[\sum_{j=1}^{r_n} I(X_j > u_n) \mid \sum_{j=1}^{r_n} I(X_j > u_n) \geq 1 \right] \longrightarrow \theta^{-1}, \quad (3)$$

with $I(X > u_n)$ the indicator function. By stationarity, this property is satisfied for any block of r_n consecutive elements defined in the sequence.

Inference about the extremal index parameter has also been extensively studied. The most popular estimators are the logs method, obtained from operating with the asymptotic results of the distribution of the maximum, the runs method derived from operating with (2) and the blocks method obtained from (3). For a careful review of these and other related estimators proposed in the literature, see [5,7,8]. Alternative estimation techniques recently developed are [9–12]. Finally, for a review of the underlying probability theory, the interested reader can consult [13] or, more recently, [14].

Our first aim in this paper is to build on the results of Leadbetter given in (3) about cluster size and to introduce an alternative characterization of θ as a limiting probability characterized by two Poisson processes. This characterization allows an intuitive and simple estimation procedure. Instead of focusing on the cluster size of extremes in a sequence, our characterizing condition of θ focuses on an extra level v_n satisfying the following property

$$E \left[\sum_{j=1}^{r_n} I(X_j > v_n) \mid \sum_{j=1}^{r_n} I(X_j > u_n) \geq 1 \right] \longrightarrow 1. \quad (4)$$

We will see that this condition implies that the ratio of exceedances of v_n and u_n by the sequence of block maxima converges asymptotically to θ .

This characterization of the extremal index as a limiting probability naturally yields a family of estimators that is consistent and converges, after proper standardization, to a normal distribution. This result is new in the extreme value theory literature. In fact, most estimators of θ proposed in the literature are inconsistent as the sample size increases. This is due to the Poisson property inherited from the choice of extreme levels for defining the extremal index estimator. A few exceptions are the estimators of θ proposed in [9,15] and, more recently, [14]. The first author solves this problem by using lower levels that allow one to benefit from increasing sample sizes and also by introducing a standardizing sequence that corrects for increasing cluster sizes. In [15], using this sequence, a variant of the blocks method is proposed that is asymptotically normal. In a similar context, Novak and Weissman [9] also prove the consistency and asymptotic normality of the blocks and runs methods.

Robert [14] introduces estimators of the limiting cluster size probabilities, which are constructed through a recursive algorithm. This author derives estimators of the extremal index and studies their asymptotic and finite-sample properties. Other recent articles generalizing the extreme value theory to models with serial dependence are, for example, [16–18].

In this paper the characterization of θ and the subsequent estimator determined by two levels makes statistical inference about the θ parameter possible in a simpler manner. Interestingly, the estimation procedure is straightforward and not very sensitive to the practical choice of the block length and threshold u_n for fixed sample sizes. Similar strategies are pursued in [9,15] and, more recently, [14]. We also show in an extensive simulation study that the nominal coverage of our estimator obtained from the asymptotic normal distribution is very good for small sample sizes. Finally, our estimator fares very well in a comparison of finite samples among the logs, blocks and runs methods for a wide class of time series exhibiting clustering of extremes and widely discussed in the literature.

The paper is structured as follows. Section 2 discusses the new characterization of the extremal index. Section 3 introduces the family of estimators of θ and its asymptotic properties. In Section 4, we illustrate these asymptotic results with a simulation experiment for time series exhibiting clustering of extreme values. In Section 5, our alternative estimation method is applied to test for the presence of clustering of extremes in monthly macroeconomic time series of unemployment growth and inflation rates in the United States. Section 6 concludes.

2. Characterization of the Extremal Index

It is well known for sequences of independent and identically distributed (i.i.d.) random variables following an unknown distribution F that, under some regularity conditions, the asymptotic distribution of the sample maximum is non-degenerate. In particular, for some suitable constants $a_n > 0$, b_n , we have that

$$P\{a_n^{-1}(M_{1,n} - b_n) \leq x\} \longrightarrow G(x), \tag{5}$$

where G is the relevant limiting distribution function that must be one of the following types (see [19]),

Type I: (Gumbel) $G(x) = e^{-e^{-x}}, \quad -\infty < x < \infty.$

Type II: (Fréchet) $G(x) = \begin{cases} 0 & x \leq 0, \\ e^{-x^{-\xi}} & x > 0, \xi > 0. \end{cases}$

Type III: (Weibull) $G(x) = \begin{cases} 1 & x \geq 0, \\ e^{-(-x)^{-\xi}} & x < 0, \xi < 0. \end{cases}$

Taking logs in both terms of (5) and denoting $u_n(x) = a_n x + b_n$, we observe that

$$n(1 - F(u_n(x))) \longrightarrow \tau(x) \quad \text{as } n \rightarrow \infty, \tag{6}$$

with $\tau(x)$ a positive real function defined by the exponent of any of the three extreme value distributions introduced above.

For large n and u_n sufficiently high, Equation (6) is sufficient to define a family of random variables $Z_{u_n(x)} := \sum_{i=1}^n I(X_i > u_n(x))$ indexed by x that converges in distribution to a family of Poisson random

variables with mean $\tau(x)$; see [20,21]. For the sake of simplicity in the exposition, we will assume hereafter x fixed and will use u_n to denote the threshold sequence $u_n(x)$. Similarly, Z_{u_n} denotes the corresponding random variable.

These important and well-known results of extreme value theory (evt) can be extended to study the maximum of a wide class of dependent processes. We concentrate here on stationary sequences where the extent of long-range dependence is restricted by a distributional mixing condition $D(u_n)$ introduced in [3]. This mixing condition is said to hold for a sequence $\{u_n\}$ if for any integers $1 \leq i_1 < \dots < i_p < j_1 < \dots < j_{p'} \leq n$ for which $j_1 - i_p \geq l_n$, we have

$$D(u_n) : \left| F_{i_1, \dots, i_p, j_1, \dots, j_{p'}}(u_n) - F_{i_1, \dots, i_p}(u_n)F_{j_1, \dots, j_{p'}}(u_n) \right| \leq \alpha_{n, l_n}, \tag{7}$$

where $\alpha_{n, l_n} \rightarrow 0$ as $n \rightarrow \infty$ for some $l_n = o(n)$ and where $F_{i_1, \dots, i_p}(u_n)$ denotes $P\{X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n\}$. This condition entails the asymptotic serial independence of the extreme events, these defined as exceedances over the threshold u_n .

Under this mixing condition, expressions (1) and (5) guarantee that

$$P\{M_{1,n} \leq u_n\} \longrightarrow G^\theta(x), \quad \text{as } n \rightarrow \infty. \tag{8}$$

Furthermore, [3] showed that there exist different partitions of the sequence $\{X_i, i \geq 1\}$ of length n defined by k_n blocks of size r_n , with $k_n \rightarrow \infty, k_n = o(n), k_n l_n = o(n)$ with l_n , introduced in (7), $r_n = [n/k_n]$, with $[\cdot]$ the integer part, such that

$$P\{M_{1,n} \leq u_n\} - P^{k_n}\{M_{1,r_n} \leq u_n\} \longrightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{9}$$

This approximation of the asymptotic distribution of the sample maximum under serial dependence and condition (8) imply, after taking logs, that

$$k_n (1 - F_{1, \dots, r_n}(u_n)) \longrightarrow \theta\tau. \tag{10}$$

In this environment, the random variable Z_{u_n} does not consist of independent elements and, in general, no longer converges in distribution to a Poisson random variable. Nonetheless, this random variable can be thinned to eliminate the presence of serial short-range dependence in the extremes. The thinning process consists of dividing the sequence of length n in k_n blocks of size r_n and choosing the block maxima that exceed the level u_n . This method allows one to define a new random variable denoted $Z_{u_n}^* := \sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > u_n)$ whose observations, under $D(u_n)$, are asymptotically serially independent.

Theorem 4.1 of [3] uses this thinning to define a point process $N_n^{(u_n)}$ on the interval $(0, 1]$ consisting of the elements of $Z_{u_n}^*$ indexed by $j/k_n, j = 1, \dots, k_n$ and that converges in distribution to a Poisson process with mean $\theta\tau$ and denoted hereafter $N(\theta\tau)$.

We build on these results, in particular condition (3), to define a sequence of extreme levels v_n characterized by the following condition

$$E \left[\sum_{j=1}^{r_n} I(X_j > v_n) \mid \sum_{j=1}^{r_n} I(X_j > u_n) \geq 1 \right] \longrightarrow 1, \tag{11}$$

where v_n depends on the choice of u_n and satisfies by construction that $v_n \geq u_n$. Further, given that

$$E \left[\sum_{j=1}^{r_n} I(X_j > v_n) \mid \sum_{j=1}^{r_n} I(X_j > u_n) \geq 1 \right] = \frac{r_n P\{X_j > v_n\}}{P \left\{ \bigcup_{j=1}^{r_n} (X_j > u_n) \right\}}, \tag{12}$$

and satisfies (11), it follows by (10) and by multiplying the numerator and denominator on the right term by k_n that

$$n(1 - F(v_n)) \longrightarrow \theta\tau, \quad \text{with } 0 < \tau < \infty \text{ as } n \rightarrow \infty. \tag{13}$$

This condition implies that

$$P\{M_{1,n} \leq v_n\} \longrightarrow G^{\theta^2} \text{ as } n \rightarrow \infty, \tag{14}$$

and therefore, for appropriate sequences k_n and r_n , this is equivalent to

$$k_n(1 - F_{1,\dots,r_n}(v_n)) \longrightarrow \theta^2\tau \text{ as } n \rightarrow \infty. \tag{15}$$

The sequence v_n defines a further thinning of $Z_{u_n}^*$ given by $Z_{v_n}^* := \sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > v_n)$ and an associated point process $N_n^{(v_n)}$ in $(0, 1]$, indexed by j/k_n , that satisfies the following result:

Theorem 1. *Let the stationary sequence $\{X_i\}_{i=1}^n$ satisfy $D(u_n)$ where u_n satisfies (6). Let $k_n \rightarrow \infty$, $k_n = o(n)$ and $k_n l_n = o(n)$ with l_n introduced in (7). Let $\{X_i\}_{i=1}^n$ have extremal index θ , with $0 < \theta \leq 1$. Then, the point process $N_n^{(v_n)}$, with v_n satisfying (11), converges in distribution to a Poisson process $N'(\theta^2\tau)$ on $(0, 1]$.*

The proof of this result follows from Theorem 4.1 of [3] and the above conditions. This theorem allows us to introduce an alternative characterization of the extremal index as the ratio of the limiting point processes $N_n^{(v_n)}$ and $N_n^{(u_n)}$.

Corollary 1. *Under assumptions in Theorem 1, the extremal index θ is the ratio of the intensity parameters of the Poisson processes $N'(\theta^2\tau)$ and $N(\theta\tau)$, defined as the limits of the processes of $N_n^{(v_n)}$ and $N_n^{(u_n)}$, respectively.*

Alternatively, we observe that by operating with expressions (10) and (15), the extremal index θ can be characterized as in the following corollary:

Corollary 2. *Under assumptions in Theorem 1, the extremal index θ is characterized as the limit of the probability $P\{M_{1,r_n} > v_n \mid M_{1,r_n} > u_n\}$.*

The proof of this result is immediate by observing that

$$P\{M_{1,r_n} > v_n \mid M_{1,r_n} > u_n\} = \frac{1 - F_{1,\dots,r_n}(v_n)}{1 - F_{1,\dots,r_n}(u_n)} \longrightarrow \theta, \quad \text{as } k_n \rightarrow \infty. \tag{16}$$

The sequences v_n and u_n are extreme levels that determine two point processes that converge in the limit to the above Poisson processes. In contrast to the standard versions of the logs, blocks and runs

methods, the family of estimators of θ derived from the new characterizations above are consistent, and their asymptotic distribution is Gaussian. We explore these properties in the following section.

3. Estimation of the Extremal Index

The extremal index provides a measure of the clustering of the largest observations of a stationary sequence. We review in this section three of the most popular estimators of θ introduced in the literature: the logs method, the blocks method and the runs estimator. The section follows with the definition of a novel consistent and asymptotically normal estimator of θ that stems naturally from our characterization of the extremal index as the limiting probability defined by the sequences u_n and v_n .

The logs method builds on the approximation of the asymptotic distribution of $M_{1,n}$ given by $P^{k_n}\{M_{1,r_n} \leq u_n\}$, with $\{k_n, r_n\}$ an appropriate partition of the sample of size n . The estimator takes the form

$$\hat{\theta}_n^{(1)} = \frac{\log(1 - Z_{u_n}^*/k_n)}{r_n \log(1 - Z_{u_n}/n)}, \tag{17}$$

with Z_{u_n}/n the empirical counterpart of $1 - F(u_n)$ and $Z_{u_n}^*/k_n$ of $1 - F_{1,\dots,r_n}(u_n)$.

Alternatively, the concept of extremal index introduced by [3] and given by interpreting θ^{-1} as the limiting mean cluster size of the exceedances yields the blocks method:

$$\hat{\theta}_n^{(2)} = \frac{Z_{u_n}^*}{Z_{u_n}}. \tag{18}$$

This estimator can be regarded as an approximation of $\hat{\theta}_n^{(1)}$ using the first order expansions of the logarithm for the numerator and denominator. Another popular estimator is the runs estimator; this method naturally follows from the characterization of θ in [6] and is given by

$$\bar{\theta}_n = \frac{W_{u_n}}{Z_{u_n}}, \tag{19}$$

where $W_{u_n} = \sum_{i=1}^{n-r_n} I(X_i > u_n)(1 - I(X_{i+1} > u_n)) \cdot \dots \cdot (1 - I(X_{i+r_n} > u_n))$.

The first two statistical moments of these estimators are studied in [4]. However, due to the Poisson character of $Z_{u_n}^*$ entailed by the choice of extreme levels u_n and the potential strong dependence between adjacent observations in Z_{u_n} , the consistency and asymptotic distribution of these estimators are compromised and only achieved after cumbersome standardizations; see [15], for example. In what follows, we introduce a new family of estimators for θ indexed by the sequence r_n obtained from the partition of n in k_n blocks and based on the novel characterization of the extremal index in Corollary 2. Furthermore, under appropriate choices of k_n satisfying conditions stated in Theorem 1, this family of estimators is consistent and asymptotically normal.

Definition 1. Let u_n and v_n be two sequences satisfying $D(\cdot)$, property (11) and the conditions in Theorem 1. Then, we can define the following family of estimators of θ indexed by r_n as

$$\tilde{\theta}_n(r_n) = \frac{Z_{v_n}^*}{Z_{u_n}^*}. \tag{20}$$

This estimator can be interpreted as a refinement of the blocks method in which the sequence v_n satisfies the empirical version of condition (11), that is $Z_{u_n}^* = Z_{v_n}$, and replaces u_n in (18). This specific choice of the sequence v_n with $v_n > u_n$ implies that our estimator, in contrast to the blocks method, consists of a ratio of asymptotically i.i.d. processes, and as we will see in the result below, it is consistent and asymptotically normal. Alternative refinements of the blocks method are given in [9].

Before introducing the consistency and asymptotic normality of $\tilde{\theta}_n(r_n)$, we need the following corollary:

Corollary 3. Let $Z_{u_n}^*$ and $Z_{v_n}^*$ be quantities defined by sequences u_n and v_n that satisfy $D(\cdot)$, property (6) and a partition determined by the sequence k_n satisfying the conditions in Theorem 1. Then:

$$Z_{u_n}^* = k_n(1 - F_{1,\dots,r_n}(u_n)) + o_P(k_n), \quad \text{with } k_n \rightarrow \infty \tag{21}$$

and

$$Z_{v_n}^* = k_n(1 - F_{1,\dots,r_n}(v_n)) + o_P(k_n), \quad \text{with } k_n \rightarrow \infty. \tag{22}$$

Proof of Corollary 3. For the sake of space, we only show here the proof for the sequence u_n . This result can be obtained from applying Chebyshev’s inequality to the quantity $\frac{1}{k_n} \sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > u_n)$. More specifically, from this inequality and given that $I(M_{(j-1)r_n+1, jr_n} > u_n)$ are Bernoulli random variables with variance $(1 - F_{1,\dots,r_n}(u_n))F_{1,\dots,r_n}(u_n)$, we know that

$$P\left\{\left|\frac{1}{k_n} \sum_{j=1}^{k_n} I_j - (1 - F_{1,\dots,r_n}(u_n))\right| > \varepsilon\right\} \leq \frac{(1 - F_{1,\dots,r_n}(u_n))F_{1,\dots,r_n}(u_n)}{k_n \varepsilon^2} + \frac{2 \sum_{i=1}^{k_n} \sum_{j>i}^{k_n} cov(I_i, I_j)}{k_n^2 \varepsilon^2}, \tag{23}$$

where, for the sake of space in the expression, $I_j := I(M_{(j-1)r_n+1, jr_n} > u_n)$.

By definition of the sequence u_n and condition $D(u_n)$, the right term on the preceding expression converges to zero, for $k_n \rightarrow \infty$ and for every $\varepsilon > 0$. Then

$$Z_{u_n}^* := \sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > u_n) = k_n(1 - F_{1,\dots,r_n}(u_n)) + o_P(k_n). \tag{24}$$

Further, the Poisson character of the quantity $\sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > u_n)$ and (10) imply that

$$Z_{u_n}^* = \theta\tau + o_P(k_n), \quad \text{for } k_n \rightarrow \infty, \tag{25}$$

with $0 < \tau < \infty$ constant. \square

With these results in place, we are ready to introduce the following results.

Theorem 2. Let u_n and v_n be sequences satisfying $D(\cdot)$, property (11) and the conditions in Theorem 1. Then

$$\tilde{\theta}_n(r_n) = \theta + o_P(1), \quad \text{as } k_n \rightarrow \infty, \tag{26}$$

and

$$\sqrt{k_n(1 - F_{1,\dots,r_n}(u_n))} \left(\tilde{\theta}_n(r_n) - \theta\right) \xrightarrow{d} N(0, \theta), \quad \text{as } k_n \rightarrow \infty, \tag{27}$$

with d denoting convergence in the distribution.

Proof of Theorem 2. Let u_n and v_n be sequences satisfying $D(u_n)$, $D(v_n)$, (11) and the conditions in Theorem 1. Let $\tilde{\theta}_n(r_n)$ be the family of estimators of θ introduced above and $\theta_n = \frac{1-F_{1,\dots,r_n}(v_n)}{1-F_{1,\dots,r_n}(u_n)}$. Now,

by Chebyshev’s inequality applied to the quantity $\frac{1}{k_n} \frac{\sum_{j=1}^{k_n} I(M_{(j-1)r_n+1,jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}}$, we obtain

$$P\left\{\left|\frac{1}{k_n} \frac{\sum_{j=1}^{k_n} I(M_{(j-1)r_n+1,jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} - \frac{(1-F_{1,\dots,r_n}(v_n))}{\sqrt{1-F_{1,\dots,r_n}(u_n)}}\right| > \varepsilon\right\} \leq \tag{28}$$

$$\leq \frac{(1-F_{1,\dots,r_n}(v_n))F_{1,\dots,r_n}(v_n)}{k_n(1-F_{1,\dots,r_n}(u_n))\varepsilon^2} + \frac{2\sum_{i=1}^{k_n}\sum_{j>i}^{k_n} cov(I(M_{(i-1)r_n+1,ir_n} > v_n), I(M_{(j-1)r_n+1,jr_n} > v_n))}{k_n^2\varepsilon^2}, \tag{29}$$

that for every $\varepsilon > 0$ converges to zero, by condition $D(v_n)$ and given that $F_{1,\dots,r_n}(v_n) \rightarrow 1$, as $k_n \rightarrow \infty$. Therefore, it holds that

$$\frac{\sum_{j=1}^{k_n} I(M_{(j-1)r_n+1,jr_n} > v_n)}{k_n\sqrt{1-F_{1,\dots,r_n}(u_n)}} = \frac{(1-F_{1,\dots,r_n}(v_n))}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} + o_P(1). \tag{30}$$

Now, by Corollary 3,

$$\frac{Z_{u_n}^*}{k_n} = 1 - F_{1,\dots,r_n}(u_n) + o_P(1). \tag{31}$$

The ratio of the last two expressions implies that

$$\frac{\tilde{\theta}_n(r_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} = \frac{\theta_n}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} + o_P(1). \tag{32}$$

Finally, from Corollary 2, it follows that $\theta_n = \theta + o(1)$ as $k_n \rightarrow \infty$. Then

$$\left(\tilde{\theta}_n(r_n) - \theta\right) = o_P\left(\sqrt{1-F_{1,\dots,r_n}(u_n)}\right). \tag{33}$$

For notational convenience, we can write the former result as

$$\sqrt{1-F_{1,\dots,r_n}(u_n)}\left(\tilde{\theta}_n(r_n) - \theta_n\right) = o_P(1-F_{1,\dots,r_n}(u_n)),$$

that implies

$$\sqrt{1-F_{1,\dots,r_n}(u_n)}\left(\tilde{\theta}_n(r_n) - \theta_n\right) = o_P(1),$$

given that $1 - F_{1,\dots,r_n}(u_n) \rightarrow 0$ as $k_n \rightarrow \infty$.

For the second result of the theorem, we first note that expression (30) multiplied by the normalizing rate $\sqrt{k_n}$ can be written as

$$\frac{1}{\sqrt{k_n}} \sum_{j=1}^{k_n} \left(\frac{I(M_{(j-1)r_n+1,jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} - \frac{1-F_{1,\dots,r_n}(v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} \right). \tag{34}$$

Now, by condition $D(u_n)$, the observations $\left(\frac{I(M_{(j-1)r_n+1, jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} - \frac{1-F_{1,\dots,r_n}(v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} \right)$ have zero mean and are asymptotically serially independent. Furthermore, the variance of each observation is given by

$$V \left(\frac{I(M_{(j-1)r_n+1, jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} \right) = \frac{(1-F_{1,\dots,r_n}(v_n))F_{1,\dots,r_n}(v_n)}{1-F_{1,\dots,r_n}(u_n)} = \theta_n F_{1,\dots,r_n}(v_n). \tag{35}$$

The observations are indexed by the block size r_n corresponding to the partition k_n . This implies that the standard central limit theorem results cannot be applied. Instead, we have to use the triangular array version of the Lindeberg-Lévy central limit theorem defined by elements $Y_{k_n j} = \frac{I(M_{(j-1)r_n+1, jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}}$. The Lindeberg condition in this framework is

$$\lim_{k_n \rightarrow \infty} \sum_{j=1}^{k_n} (s_{k_n}^{-2}) E [Y_{k_n j}^2 I(|Y_{k_n j}| \geq \varepsilon s_{k_n})] = 0, \tag{36}$$

with ε any positive value and where $s_{k_n}^2 = \sum_{j=1}^{k_n} V(Y_{k_n j}) + 2 \sum_{i=1}^{k_n} \sum_{j>i}^{k_n} cov(Y_{k_n i}, Y_{k_n j})$ with $V(Y_{k_n j})$ obtained in (35) and $cov(Y_{k_n i}, Y_{k_n j}) \rightarrow 0$ as $k_n \rightarrow \infty$, by the mixing condition $D(v_n)$. After some basic algebra, it is simple to see that (36) is satisfied, and then, as $k_n \rightarrow \infty$,

$$\frac{1}{\sqrt{k_n}} \sum_{j=1}^{k_n} \left(\frac{I(M_{(j-1)r_n+1, jr_n} > v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} - \frac{1-F_{1,\dots,r_n}(v_n)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} \right) \xrightarrow{d} N \left(0, \lim_{k_n \rightarrow \infty} \theta_n F_{1,\dots,r_n}(v_n) \right). \tag{37}$$

Now, by the consistency result (31) and dividing by $\frac{Z_{u_n}^*}{k_n}$, we obtain

$$\sqrt{k_n(1-F_{1,\dots,r_n}(u_n))} \left(\tilde{\theta}_n(r_n) - \theta_n \right) \xrightarrow{d} N \left(0, \lim_{k_n \rightarrow \infty} \theta_n F_{1,\dots,r_n}(v_n) \right). \tag{38}$$

Finally, the choice of v_n implies that $F_{1,\dots,r_n}(v_n) \rightarrow 1$, and we obtain the desired result. \square

In fact, the true rate of convergence for the consistency is $\frac{1}{\sqrt{1-F_{1,\dots,r_n}(u_n)}}$. This is so because

$$\frac{\left(\tilde{\theta}_n(r_n) - \theta \right)}{\sqrt{1-F_{1,\dots,r_n}(u_n)}} = o_P(1), \quad \text{as } k_n \rightarrow \infty,$$

and $1-F_{1,\dots,r_n}(u_n) \rightarrow 0$; see the first part of the proof of Theorem 2 for details. This interesting result implies that

$$\sqrt{1-F_{1,\dots,r_n}(u_n)} \left(\tilde{\theta}_n(r_n) - \theta \right) = o_P(1), \quad \text{as } k_n \rightarrow \infty.$$

Hence, from (27), we achieve, as in standard statistical theory, a $\sqrt{k_n}$ rate of convergence for the asymptotic normality.

It is still unresolved how to choose in practice the sequence v_n conditional on an appropriate choice of the sequence u_n satisfying $D(u_n)$ and determining $Z_{u_n}^*$. A suitable choice for v_n is the sequence that satisfies the empirical counterpart of condition (11);

$$\frac{1}{Z_{u_n}^*} \sum_{j=1}^{k_n} \sum_{i=(j-1)r_n+1}^{jr_n} I(X_i > \hat{v}_n) \rightarrow 1,$$

that yields, naturally, the following estimator of v_n :

$$\hat{v}_n = F^{-1} \left(1 - \frac{Z_{u_n}^*}{n} \right). \tag{39}$$

With this sequence, we define the process $Z_{\hat{v}_n}^* = \sum_{j=1}^{k_n} I(M_{(j-1)r_n+1, jr_n} > \hat{v}_n)$ and obtain the corresponding feasible version of the estimator $\tilde{\theta}_n(r_n)$, defined now by

$$\tilde{\theta}_n^f(r_n) = \frac{Z_{\hat{v}_n}^*}{Z_{u_n}^*}.$$

Furthermore, by the consistency result obtained in Corollary 3, we can replace the rate $\sqrt{k_n(1 - F_{1, \dots, r_n}(u_n))}$ by $\sqrt{Z_{u_n}^*}$ and obtain the following statistic:

$$\sqrt{Z_{u_n}^*} \left(\tilde{\theta}_n^f(r_n) - \theta \right).$$

Corollary 4. Let u_n be a sequence satisfying $D(u_n)$ and defining the quantity $Z_{u_n}^*$, and let $\hat{v}_n = F^{-1} \left(1 - \frac{Z_{u_n}^*}{n} \right)$, with $F(x)$ the distribution function of a sequence $\{X_i\}_{i=1}^n$ with extremal index $0 < \theta \leq 1$. Furthermore, let $k_n \rightarrow \infty$, $k_n = o(n)$ and $k_n l_n = o(n)$ with l_n introduced in (7). Then

$$\tilde{\theta}_n^f(r_n) = \theta + o_P(1), \quad \text{as } k_n \rightarrow \infty, \tag{40}$$

and

$$\sqrt{Z_{u_n}^*} \left(\tilde{\theta}_n^f(r_n) - \theta \right) \xrightarrow{d} N(0, \theta), \quad \text{as } k_n \rightarrow \infty. \tag{41}$$

Proof of Corollary 4. We will concentrate on the consistency of the estimator. The proof of asymptotic normality, once the consistency result is derived follows immediately from dividing and multiplying $\sqrt{k_n}$ by $\sqrt{Z_{u_n}^*}$ in (41), from the consistency result (31) and from applying Theorem 2.

Let u_n be a sequence satisfying $D(u_n)$ and defining the quantity $Z_{u_n}^*$, and let $\hat{v}_n = F^{-1} \left(1 - \frac{Z_{u_n}^*}{n} \right)$. We first need to show that $D(\hat{v}_n)$ holds given that $D(u_n)$ does. By Lemma 3.6.2 (iv) in [22], $D(\hat{v}_n)$ will hold if $\hat{v}_n \geq u_n$. Therefore, we need to prove this inequality. A sufficient condition is to see that $F(u_n) \leq F(\hat{v}_n)$ for any given n sufficiently high. Note that $F(u_n) = 1 - \frac{Z_{u_n}^*}{n} + o_P(1)$ with $Z_{u_n} = \sum_{i=1}^n I(X_i > u_n)$. Then, $F(u_n) \leq F(\hat{v}_n)$ if and only if $Z_{u_n} \geq Z_{u_n}^*$. Now, by the character of these two quantities, $Z_{u_n}^*$ a thinning of Z_{u_n} , this inequality holds naturally.

For the second part of the consistency proof, we note that by (6) and (10), the mixing condition $D(\hat{v}_n)$ implies that

$$\frac{k_n(1 - F_{1, \dots, r_n}(\hat{v}_n))}{n(1 - F(\hat{v}_n))} \rightarrow \theta, \quad \text{as } k_n, n \rightarrow \infty. \tag{42}$$

Using this asymptotic relationship, and after some algebra, the estimator \hat{v}_n satisfies

$$\hat{v}_n = F_{1, \dots, r_n}^{-1} \left(1 - \frac{\theta n(1 - F(\hat{v}_n))}{k_n} + o_P \left(\frac{1}{k_n} \right) \right). \tag{43}$$

Now, replacing (43) into (22) and observing that $n(1 - F(\hat{v}_n)) = Z_{u_n}^*$, we obtain

$$Z_{\hat{v}_n}^* = \theta Z_{u_n}^* + o_P(1).$$

Finally, by definition of the feasible estimator $\tilde{\theta}_n^f(r_n)$, it follows that

$$\tilde{\theta}_n^f(r_n) = \theta + o_P(1).$$

□

In practice, when the distribution F is unknown, the estimator of v_n is replaced by the empirical version of $F^{-1}\left(1 - \frac{Z_{u_n}^*}{n}\right)$, that is by the extreme order statistic $X_{Z_{u_n}^*+1:n}$ from the sequence $X_{1:n} \geq X_{2:n} \geq \dots \geq X_{n:n}$. Furthermore, in the studies in finite samples with n fixed, we will use as a candidate for u_n an extreme order statistic of the sequence of block maxima. In particular, this will be $\hat{u}_n = M_{t+1:k_n}$ with $t > 0$ fixed, implying that $Z_{\hat{u}_n}^* = t$ and $\hat{v}_n = X_{t+1:n}$. These results enable us to apply Theorem 2 to make statistical inference about the extremal index. An interesting case is testing for the existence of the clustering of extremes in stationary sequences. It is well known that under limiting serial independence in the extremes, θ takes the value one. Therefore a valid test is $H_0 : \theta = 1$ against the one-sided alternative $H_A : \theta < 1$. This possibility is explored in the application to macroeconomic time series.

4. Simulations: Some Examples

This section studies some of the most popular examples of time series exhibiting clustering of extremes. We discuss in detail the following four processes: the autoregressive model of [23], the doubly-stochastic process studied in [4], the moving-maximum process with Fréchet marginals and the maximum of a lagged autoregressive process introduced by [24]. The simulation exercise is divided into two components. First, we compare estimation methods across different partitions, and second, we study the asymptotic coverage of the limiting normal distribution introduced in Theorem 2 for different sample sizes.

The first example is due to [23]. Let $\{X_i\}$ be a strictly-stationary first order autoregressive sequence driven by $X_i = \frac{1}{r}X_{i-1} + \varepsilon_i$, with $r \geq 2$ an integer, ε_i discrete uniform random variables on $\{0, 1/r, \dots, (r-1)/r\}$ and independent of X_{i-1} . The random variable X_i has a uniform distribution on $[0, 1]$, and the extremal index is $\theta = \frac{r-1}{r}$. Figures 1 and 2 report the sequence of estimates of θ for $r = 5$ and $r = 2$, respectively. The different panels in these figures report the estimates of θ obtained from the logs method, blocks method, runs method and the estimator $\tilde{\theta}_n^f$, for different levels determined by order statistics $\hat{u}_n = M_{t+1:k_n}$ and $\hat{v}_n = X_{t+1:n}$ with $t = 1, 3, 5, 7$, and $n = 200$.

A close inspection of the charts shows that the blocks and the runs method underestimate θ as r_n increases. This is so because these estimators, by construction, have a decreasing numerator as the block size increases. On the other hand, estimates derived from the logs method are very accurate for the higher threshold sequences employed, but unfortunately, as u_n decreases, the estimator exhibits problems due to the fact that every single block contains an exceedance ($Z_{\hat{u}_n}^* = k_n$). In these cases, this estimator is not well defined. In contrast, $\tilde{\theta}_n^f$ shows reliable and stable estimates of θ across all threshold levels and partitions employed.

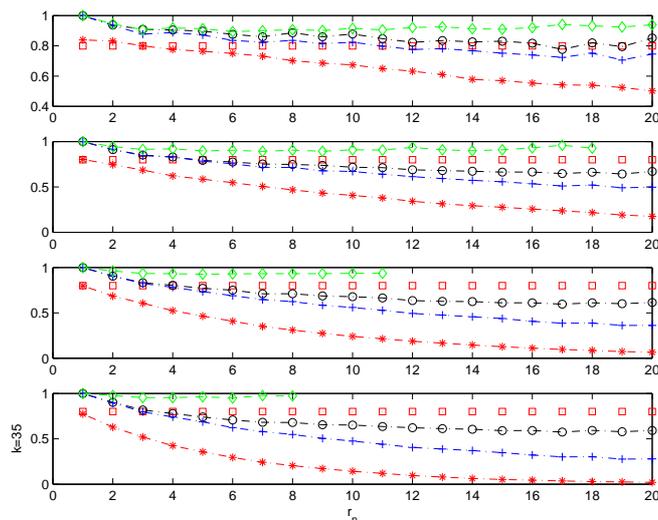


Figure 1. Chernick model with $r = 5$. $\theta = 0.8$. Sample mean for different estimators of θ . $m = 100$ Monte Carlo simulations. $\hat{u}_n = M_{t+1:k_n}$, and $\hat{v}_n = X_{t+1:n}$ with $t = 1, 3, 5, 7$; $n = 200$. $r_n \in [1, 20]$. θ plotted with, $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$, $\hat{\theta}_n^{(2)}$ with $-+$, and $\bar{\theta}_n$ with the $-*$ line.

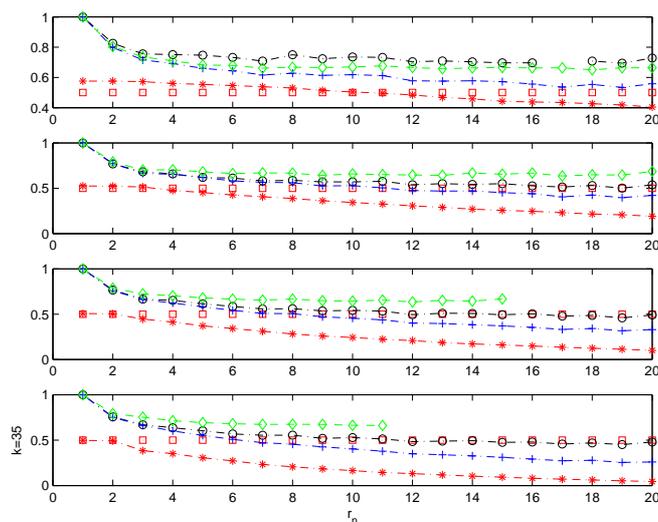


Figure 2. Chernick model with $r = 2$. $\theta = 0.5$.

The second process under investigation is the doubly-stochastic model studied in [4]. Let $\{\xi_i, i \geq 1\}$ be i.i.d. with distribution function F ; suppose that $Y_1 = \xi_1$ and for $i > 1$,

$$Y_i = \begin{cases} Y_{i-1} & \text{with probability } \psi, \\ \xi_i & \text{with probability } 1 - \psi, \end{cases}$$

the choice being made independently for each i . The doubly-stochastic sequence $\{X_i, i \geq 1\}$ is defined by

$$X_i = \begin{cases} Y_i & \text{with probability } \eta, \\ 0 & \text{with probability } 1 - \eta, \end{cases}$$

independently of anything else. In this example, the extremal index is $\theta = \frac{1-\psi}{1-\psi+\psi\eta}$. Smith and Weissman [4] compare different estimators of θ for $\psi = 0.9$ and $\eta = 0.7$ ($\theta = 0.137$) and show that the runs method is superior to the rest of the competing estimators. Figure 3 is consistent with their results. $\tilde{\theta}_n^f$ seems to be, however, a very good competitor of $\bar{\theta}_n$ for every single level and outperforms the logs and blocks estimators across all levels. To compare the performance of the runs method against $\tilde{\theta}_n^f$, we also estimate the extremal index of this process for $\psi = 0.5$ and $\eta = 0.5$ ($\theta = 0.667$). Both of the runs and blocks method exhibit the same declining pattern observed before for increasing block sizes (see Figure 4). In this case, however, $\tilde{\theta}_n^f$ shows a superior performance to the rest of the estimators. The logs methods is the only competitor exhibiting a similar performance to our family of estimators.

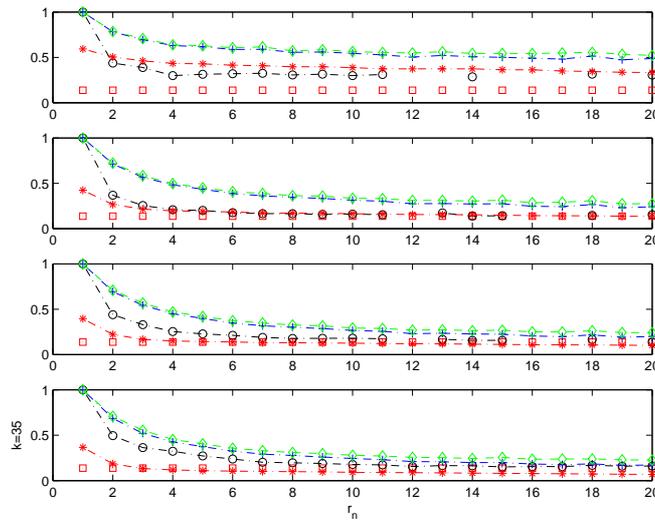


Figure 3. Doubly-stochastic model with $\psi = 0.9$ and $\eta = 0.7$. $\theta = 0.137$. Sample mean for different estimators of θ . $m = 100$ Monte Carlo simulations. $\hat{u}_n = M_{t+1:k_n}$, and $\hat{v}_n = X_{t+1:n}$, with $t = 1, 3, 5, 7$; $n = 200$. $r_n \in [1, 20]$. θ plotted with $-o$, $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$, $\hat{\theta}_n^{(2)}$ with $-+$, and $\bar{\theta}_n$ with the $-*$ line.

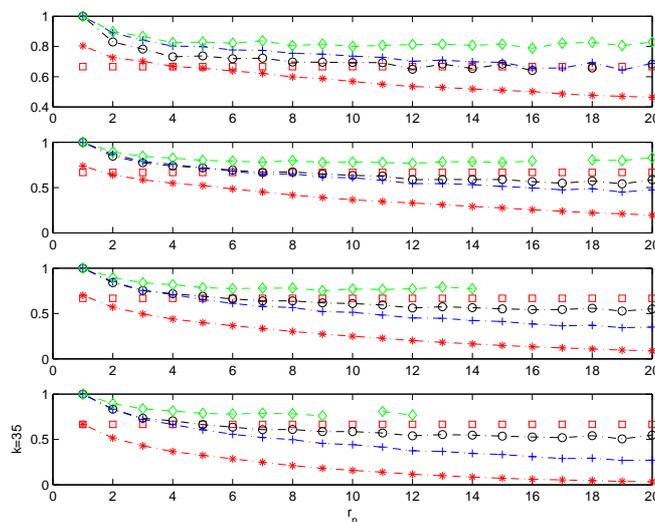


Figure 4. Doubly-stochastic model with $\psi = 0.5$ and $\eta = 0.5$. $\theta = 0.667$.

The following example is the moving-maximum process: $y_i = \max \{\mu X_i, (1 - \mu)X_{i+1}\}$, where $\{X_i, i \geq 1\}$ is an i.i.d. sequence with the common distribution function standard Fréchet ($\xi = 1$). The extremal index is $\theta = 1 - \mu$ for $\mu < 0.5$ and $\theta = \mu$, otherwise. We consider $\mu = 0.5$. Figure 5 shows results in the spirit of those found for the Chernick process with $r = 2$.

This part of the simulation section concludes with the example introduced by L. de Haan: $y_i = \max_{k \geq 0} \rho^k x_{i-k}$, where $0 < \rho < 1$ and $\{X_i, i \geq 1\}$ is an i.i.d. sequence with the common distribution function standard Fréchet ($\xi = 1$). The extremal index is $\theta = 1 - \rho$. The results reported in Figure 6 for $\rho = 0.25$ are consistent with models exhibiting low, but significant clustering; see, for example, the results of the simulation exercise for the Chernick model and $r = 5$, where $\theta = 0.8$, and the doubly-stochastic process with $\psi = 0.5$ and $\eta = 0.5$, where $\theta = 0.667$.

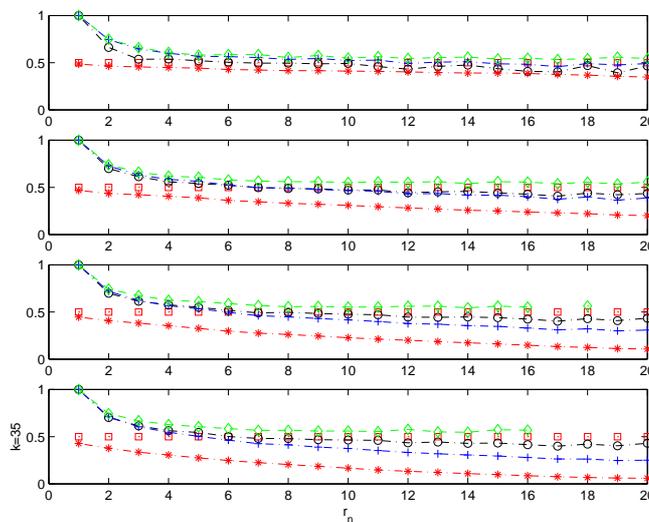


Figure 5. Moving-maximum process with $\mu = 0.5$. $\theta = 0.5$. Sample mean for different estimators of θ . $m = 100$ Monte Carlo simulations. $\hat{u}_n = M_{t+1:k_n}$, and $\hat{v}_n = X_{t+1:n}$, with $t = 1, 3, 5, 7; n = 200$. $r_n \in [1, 20]$. θ plotted with, $\hat{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$, $\hat{\theta}_n^{(2)}$ with $-+$, and $\bar{\theta}_n$ with the $-*$ line.

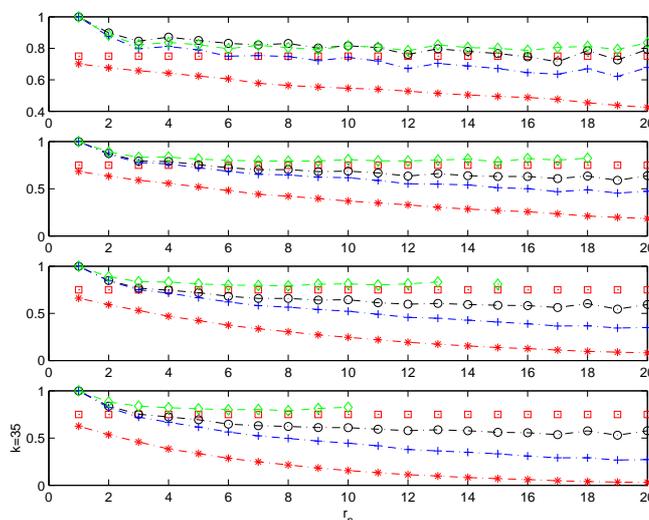


Figure 6. Maximum of the lagged autoregressive process with $\rho = 0.25$. $\theta = 0.75$.

In order to compare the results of the blocks and logs methods with our new feasible estimator more in depth, we report in Figures 7 and 8 the estimated mean square error for the doubly-stochastic model. The penalty function is

$$MSE(\tilde{\theta}_n^f(r_n)) = \frac{1}{m} \sum_{i=1}^m (\tilde{\theta}_{ni}^f(r_n) - \theta)^2, \quad \text{with } r_n = 1, \dots, 50, \tag{44}$$

with $m = 100$ simulated Monte Carlo sequences. The results confirm the findings above. Whereas the blocks method performs rather well for the first simulation experiment, it completely fails to report accurate estimates for the second experiment with a higher extremal index.

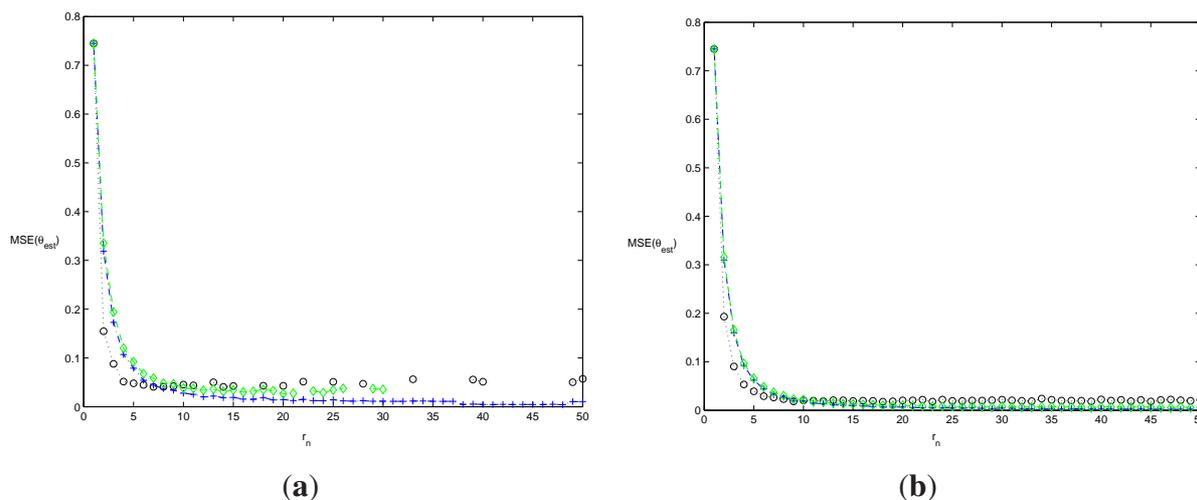


Figure 7. Simulated mean square error (MSE) of the estimators of θ for the doubly-stochastic model with $\psi = 0.9$ and $\eta = 0.7$. $m = 100$ Monte Carlo simulations. $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$ and $\hat{\theta}_n^{(2)}$ with $-+$. $r_n \in [1, 50]$. $n = 200$ in (a), and $n = 1000$ in (b). $\hat{u}_n = M_{t+1:k_n}$ and $\hat{v}_n = X_{t+1:n}$ with $t = 7$.

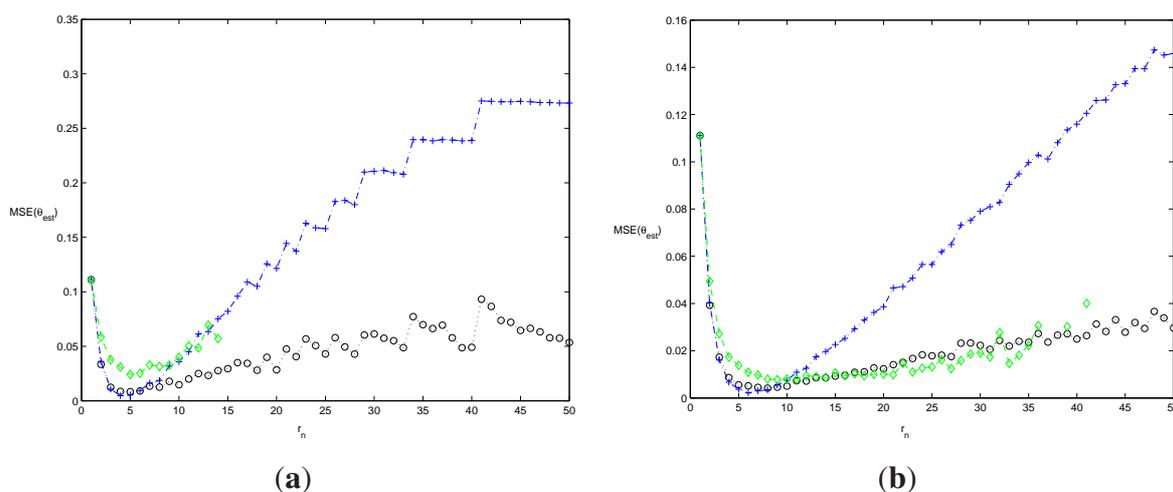


Figure 8. Simulated mean square error (MSE) of the estimators of θ for the doubly-stochastic model with $\psi = 0.5$ and $\eta = 0.5$. $m = 100$ Monte Carlo simulations. $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$ and $\hat{\theta}_n^{(2)}$ with $-+$. $r_n \in [1, 50]$. $n = 200$ in (a), and $n = 1000$ in (b). $\hat{u}_n = M_{t+1:k_n}$ and $\hat{v}_n = X_{t+1:n}$ with $t = 7$.

This section concludes with the study of the empirical coverage of the normal distribution followed asymptotically by our family of estimators. Figures 9 and 10 report the asymptotic coverage of the two-sided normal confidence intervals for the examples above covering the Chernick and doubly-stochastic models. The simulations are computed for the order statistics above with $r_n = [1, 50]$ and $n = 200, 500, 1000$. The plots illustrate the accuracy of the normal asymptotic distribution in all cases. It is worth noting that the partition of the sample and the extent of clustering play an important role for obtaining the correct nominal coverage for the asymptotic normal approximation. Thus, for large sample sizes, the models exhibiting a low clustering of extremes, Chernick with $r = 5$ and the doubly-stochastic model with $\psi = \eta = 0.5$, only provide accurate approximations as the block size increases.

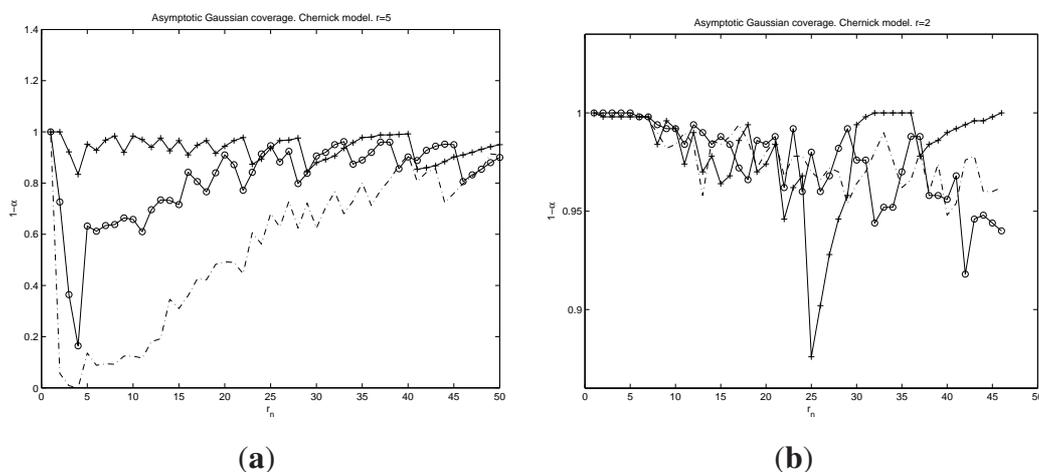


Figure 9. Chernick model with $r = 5$ in (a) and $r = 2$ in (b). $m = 500$ Monte Carlo simulations. $+ -$ for $n = 200$, $o -$ for $n = 500$ and $- -$ for $n = 1000$. $\hat{u}_n = M_{t+1:k_n}$ and $\hat{v}_n = X_{t+1:n}$ with $t = [0.005k_n]$, where $[\cdot]$ rounds to the next integer. $r_n = [1, 50]$, $k_n = n/r_n$. $\alpha = 0.05$.

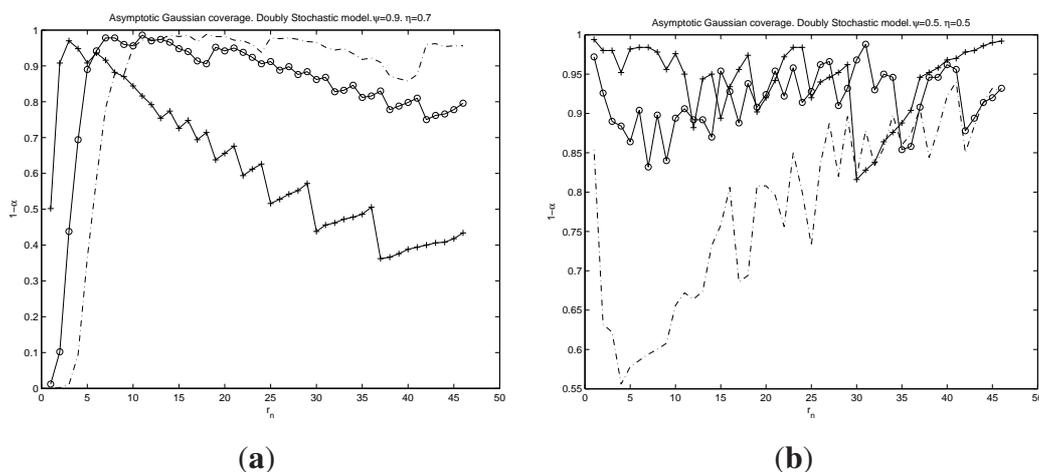


Figure 10. Doubly-stochastic model with $\psi = 0.9$ and $\eta = 0.7$ and $\psi\eta = 0.5$ in (a) and $\eta = 0.5$ in (b). $m = 500$ Monte Carlo simulations. $+ -$ for $n = 200$, $o -$ for $n = 500$ and $- -$ for $n = 1000$. $\hat{u}_n = M_{t+1:k_n}$ and $\hat{v}_n = X_{t+1:n}$ with $t = [0.005k_n]$, where $[\cdot]$ rounds to the next integer. $r_n = [1, 50]$, $k_n = n/r_n$. $\alpha = 0.05$.

5. Application to Macroeconomic Time Series

Macroeconomic series usually exhibit clustering of extreme values that indicate periods of crisis, financial distress, booms and bursts. This phenomenon is more acute as the data are studied at shorter frequencies; thus, monthly time series usually exhibit stronger clustering of extremes than quarterly series, and so on. This stylized fact in macroeconomic time series is usually modeled with econometric processes that accommodate volatility clustering; see the ARCH family of volatility processes proposed by [25] or the stochastic volatility processes developed in [26]. These processes are, however, silent about the probability of runs of extremes in one or the other tail, and only after tedious calculations, which depend most of the times on a number of parameters, can one compute the chances of these events. Alternatively, inference about the extremal index is an unexplored option in this field that can lead to very interesting insights about the existence of clustering in the extremes of these sequences and on its persistence.

In this application, we pursue this alternative with monthly data in unemployment growth and inflation rates from the United States spanning the period February 1947 until July 2007 for the first series and February 1947 until June 2008 for the second series. Data are obtained from <https://www.economy.com/freelunch>. The dynamics of these series are reported in Figures 11 and 12. Dickey-Fuller unit root tests reject the null hypothesis, providing statistical support to the stationarity of both series. As pointed out by a referee, the visual inspection of both processes could also suggest the presence of several structural breaks around periods characterized by drastic changes in monetary policy. In this likely scenario, it would also be reasonable to assume that the unemployment growth process and inflation rates are both stationary around a shifting mean, implying that the analysis of the serial dependence of the extremes of both sequences should be carried out separately for each regime. Nevertheless, for simplicity in the description of the problem, we assume hereafter a constant mean in both cases and, in turn, the stationarity of both processes.

Panel (a) of Figures 13 and 14 report the estimates of θ using the four methods discussed in the paper. We can observe from the plots a similar decaying pattern for the four families of estimates. Note, however, that, whereas the sequences of the logs, blocks and runs estimates decay as the block size r_n increases, the estimator $\tilde{\theta}_n^f$ proposed in this paper stabilizes after the first partitions around a value of 0.3 for unemployment growth and 0.55 for the inflation rate, respectively. The statistical significance of these results is carried out by computing confidence intervals for θ at 5% using the results of Theorem 2 above. These intervals are displayed on the right panels of Figures 13 and 14.

The results on the extremal index point towards a strong clustering in the positive extremes of both sequences. A more detailed analysis obtained by inspecting the confidence levels of each sequence of estimates indicates a stronger clustering for the unemployment series than for inflation.¹ This implies that periods of high unemployment are more persistent than highly inflationary periods. A simple version of the Philips curve (see [27]) shows that inflation and economic growth are positively correlated or, similarly, that inflation and unemployment are negatively correlated. These empirical

¹ A more formal alternative is to compute confidence intervals for the difference between the extremal indexes of each macroeconomic series.

observations would imply that high inflation is followed by economic growth and, hence, falls in unemployment. Our empirical analysis of the extremes of both series adds further insights into this relationship. More specifically, we observe that highly inflationary periods are less persistent than periods with large unemployment growth, suggesting that economic policies focused on producing inflation to boost economic growth can only be successful in the very short term, as inflation quickly returns to normal levels. In contrast, large unemployment growth has lingering effects that are not easily reduced by policies aimed at rising inflation.

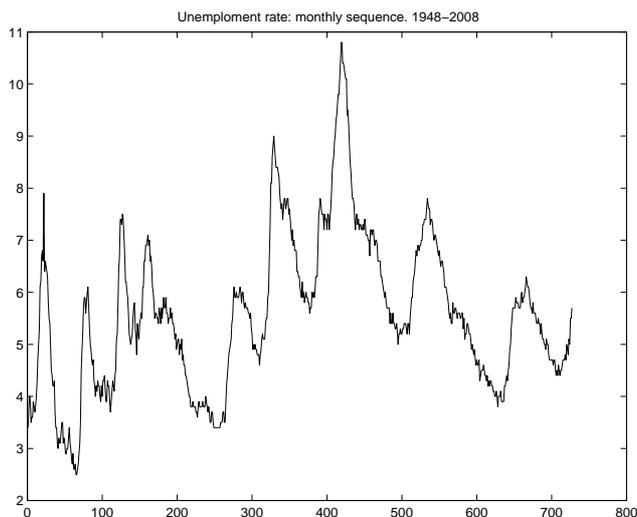


Figure 11. Monthly time series of unemployment growth for United States spanning the period February 1947 to June 2007.

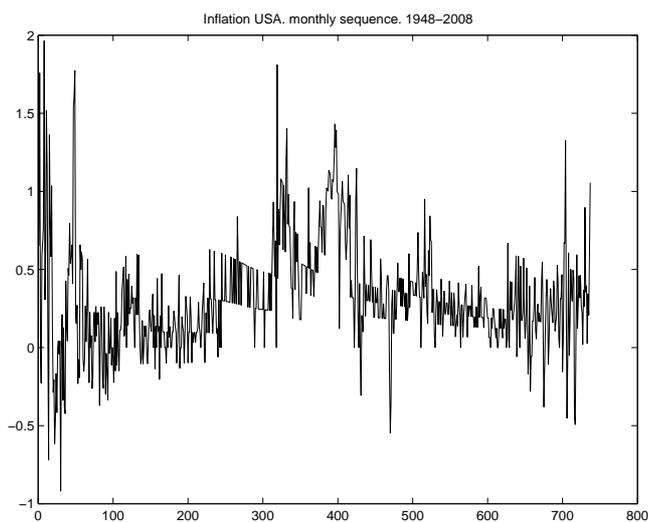


Figure 12. Monthly time series of inflation rates for the United States spanning the period February 1947 to July 2008.

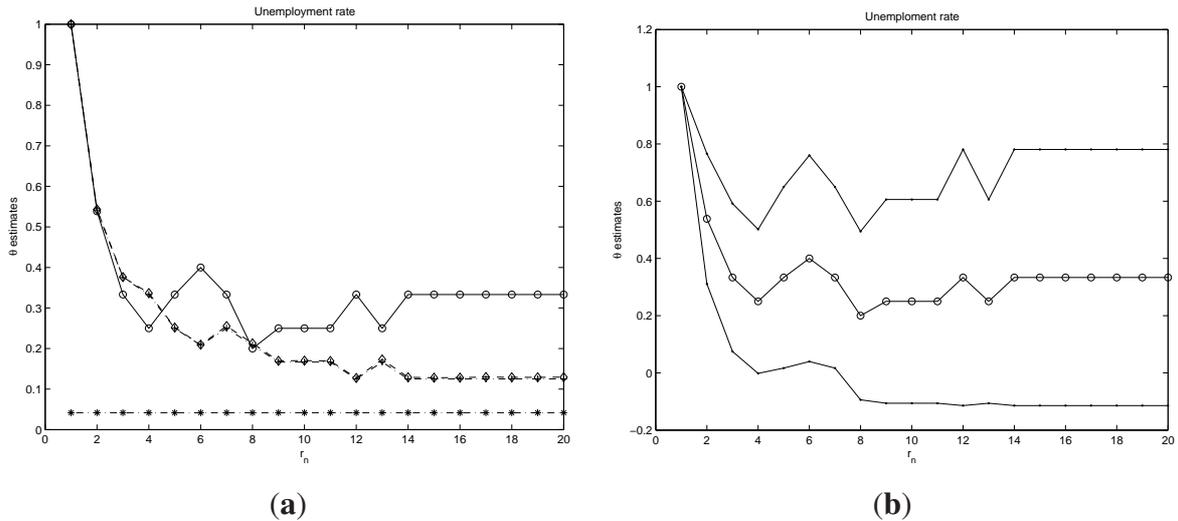


Figure 13. Panel (a) reports different estimates of θ . $r_n \in [1, 20]$. $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$, $\hat{\theta}_n^{(2)}$ with $-+$, and $\bar{\theta}_n$ with the $-*$ line. Panel (b) reports two-sided confidence intervals for θ at 5% computed from $\tilde{\theta}_n^f$, with $\hat{u}_n := M_{t+1:k_n}$ and $\hat{v}_n := X_{t+1:n}$, $t = 7$.

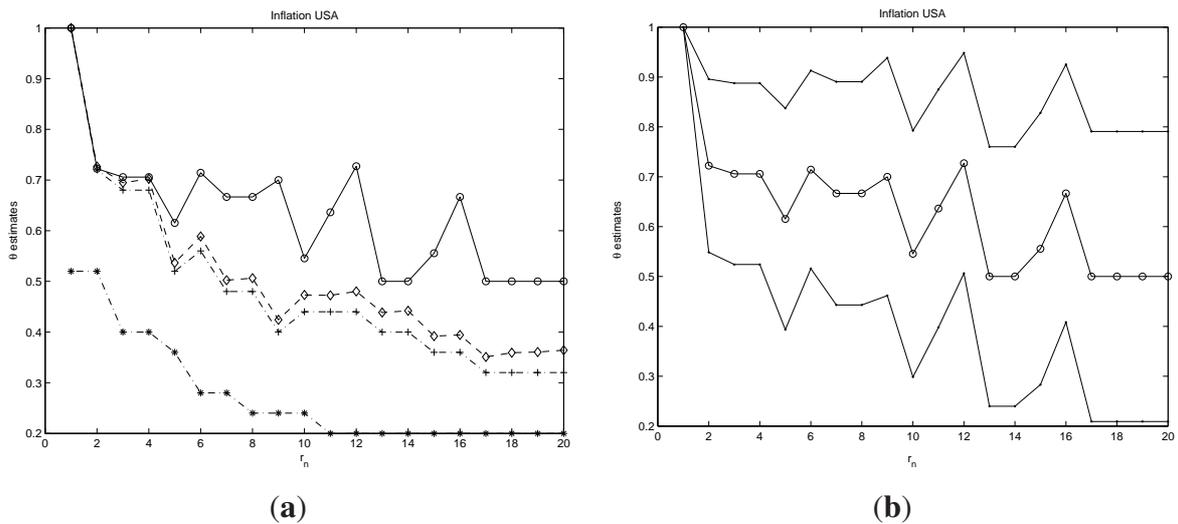


Figure 14. Panel (a) reports different estimates of θ . $r_n \in [1, 20]$. $\tilde{\theta}_n^f$ with $-o$, $\hat{\theta}_n^{(1)}$ with $-\diamond$, $\hat{\theta}_n^{(2)}$ with $-+$, and $\bar{\theta}_n$ with the $-*$ line. Panel (b) reports two-sided confidence intervals for θ at 5% computed from $\tilde{\theta}_n^f$, with $\hat{u}_n := M_{t+1:k_n}$ and $\hat{v}_n := X_{t+1:n}$, $t = 7$.

6. Conclusions

The existence of serial dependence in the extremes of stationary sequences is summarized in one single parameter: the extremal index. We have introduced in this paper a novel characterization of the extremal index based on the limiting expected value of two point processes defined by the sequence of block maxima determined by appropriate partitions and exceeding different threshold levels. Unlike most estimators for θ in the literature, this characterization yields a family of estimators that is consistent and asymptotically normal. Therefore, in contrast to most estimators of the extremal index that are downward

biased, as the block size, our family of estimators yields very stable estimates across partitions, adding reliability to the results obtained.

In our application, we observe a significant clustering of extremes in unemployment and inflation rates. A more detailed analysis to the estimates of the extremal index shows that runs of large unemployment growth are more prolonged than runs of high inflation. This indicates a higher persistence of large unemployment rates than of inflationary periods.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Loynes, R.M. Extreme Values in Uniformly Mixing Stationary Stochastic Processes. *Ann. Math. Stat.* **1965**, *36*, 993–999.
2. O'Brien, G.L. The maximum term of uniformly mixing stationary sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **1974**, *30*, 57–63.
3. Leadbetter, M. R. Extremes and Local Dependence in Stationary Sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **1983**, *65*, 291–306.
4. Smith, R.L.; Weissman, I. Estimating the Extremal index. *J. R. Stat. Soc. Ser. B* **1994**, *56*, 515–528.
5. Embrechts, P.; Klüppelberg, C.; Mikosch, T. *Modelling Extremal Events for Insurance and Finance*; Springer: Berlin, Germany, 1997.
6. O'Brien, G.L. Extreme Values for Stationary and Markov Sequences. *Ann. Probab.* **1987**, *15*, 281–291.
7. Coles, S. *An Introduction to Statistical Modeling of Extreme Values*; Springer Series in Statistics; Springer-Verlag: London, UK, 2001.
8. Finkenstadt, C.; Rootzén, H. Extreme Values in Finance, Telecommunications and Environment; In *Monographs on Statistics and Applied Probability* 99; Chapman and Hall: London, UK, 2003.
9. Novak, S.; Weissman, I. On blocks and runs estimators of the extremal index. *J. Stat. Plan. Inference* **1998**, *66*, 281–288.
10. Hsing, T. Extremal Index Estimation for a Weakly Dependent Stationary Sequence. *Ann. Stat.* **1993**, *21*, 2043–2071.
11. Ferro, C.A.; Segers, J. Inference for clusters of extremes. *J. R. Stat. Soc. Ser. B* **2003**, *65*, 545–556.
12. Laurini, F.; Tawn, J.A. New Estimators for the Extremal Index and Other Cluster Characteristics. *Extremes* **2003**, *3*, 189–211.
13. Leadbetter, M.R.; Rootzén, H. Extremal theory for stochastic processes. *Ann. Probab.* **1988**, *16*, 431–478.
14. Robert, C.Y. Inference for the Limiting Cluster Size Distribution of Extreme Values. *Ann. Stat.* **2009**, *37*, 271–310.
15. Hsing, T. Estimating the parameters of rare events. *Stoch. Process. Appl.* **1991**, *37*, 117–139.
16. Drees, H. Weighted approximations of tail processes for β -mixing random variables. *Ann. Appl. Probab.* **2000**, *10*, 1274–1301.

17. Drees, H. Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **2003**, *9*, 617–657.
18. De Haan, L.; Mercadier, C.; Zhou, C. Erasmus University of Rotterdam, Rotterdam, Netherlands. Adapting extreme value statistics to financial time series: Dealing with bias and serial dependence. Unpublished manuscript, 2015.
19. Gnedenko, B.V. Sur la distribution limit du terme d'une série aléatoire. *Ann. Math.* **1943**, *44*, 423–453.
20. Hodges, J.L.; le Cam, L. The Poisson approximation to the Poisson binomial distribution. *Ann. Math. Stat.* **1960**, *31*, 737–740.
21. Lehman, E.L. *Elements of Large-Sample Theory*; Springer-Verlag: New York, NY, USA, 1999.
22. Leadbetter, M.R.; Lindgren, G.; Rootzén, H. *Extremes and Related Properties of Random Sequences and Processes*; Springer-Verlag: New York, NY, USA, 1983.
23. Chernick, M.R. A limit theorem for the maximum of autorregressive processes with uniform marginal distribution. *Ann. Probab.* **1981**, *9*, 145–149.
24. De Haan, L. Sample extremes: An elementary introduction. *Stat. Neerlandica* **1976**, *30*, 161–172.
25. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007.
26. Taylor, S.J. Modeling stochastic volatility: A review and comparative study. *Math. Financ.* **1994**, *4*, 183–204.
27. Phillips, A. W. The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861–1957. *Economica* **1958**, *25*, 283–299.

© 2015 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).