Dynamic threshold modelling and the US business cycle

M. de Carvalho

Pontificia Universidad Católica de Chile, Chile
Universidade Nova de Lisboa, Portugal

K. F. Turkman

Universidade de Lisboa, Portugal

A. Rua

Banco de Portugal, Portugal

Abstract. Leading economic indicators are often used to anticipate changes in key economic variables. Understanding the dynamics of these indicators is of primary interest for policy-making objectives and for a sustainable economic welfare. In this paper we are concerned with the problem of setting a dynamic threshold above which the value of such leading indicator would be considered as extreme. We propose a dynamic threshold modelling approach based on fractionally integrated processes where a semi-parametric method is used to determine the amount of differencing required to obtain a weakly stationary process—to which standard methods of statistics of extremes apply. Given that our approach is linked to the Box–Jenkins method, we refer to the procedure proposed and applied herein as Box–Jenkins–Pareto. We use our approach to analyze the weekly number of unemployment insurance claims in the US and explore the connection between its threshold exceedances and the US business cycle.

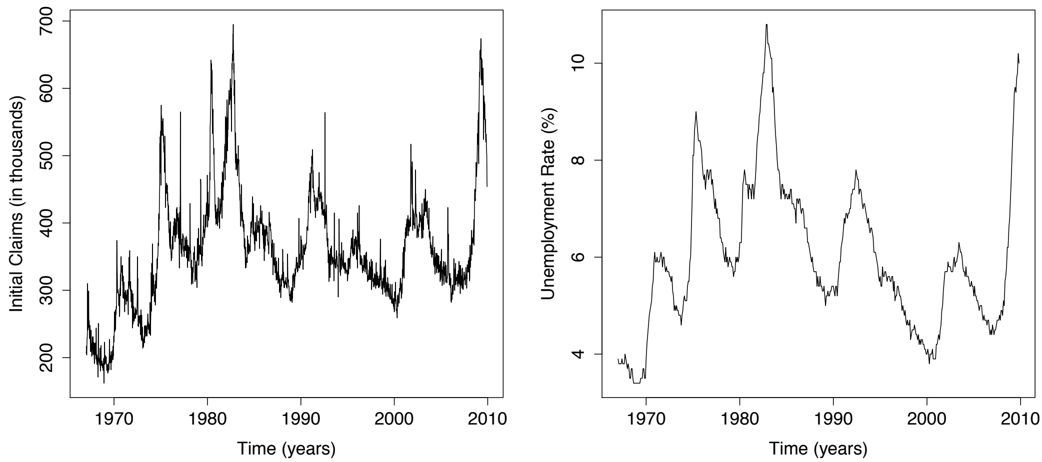
Keywords: Box–Jenkins method; Extreme value econometrics; Nonstationary process; Peaks over threshold; Statistics of extremes; Unemployment data; US business cycle

1. Introduction

The recent US subprime crisis and the current European sovereign-debt crisis, have been increasing awareness for the need to model extreme values. Among the most used techniques for modelling extremes are peaks over threshold methods Coles (2001, Ch. 4). These methods consider as extreme all observations above a fixed large threshold, which yields a Generalized Pareto Distribution (GPD) as the limiting distribution of the threshold exceedances (Balkema and de Haan, 1974; Pickands, 1975). Such methods are developed for weakly stationary time series $\{Y_t\}_{t\in\mathbb{Z}}$ that by definition fluctuate around a fixed level $\mathrm{E}(Y_t) = \mu$, and whose dependence on past observations is such that there exists a real-valued function $g(|t-\ell|) = \mathrm{cov}(Y_t, Y_\ell)$, for every $\ell \in \mathbb{Z}$; below the word 'stationary' will be used to abbreviate 'weakly stationary' as defined here. Many data of interest fail however to be stationary, and in this paper we provide a dynamic threshold modelling approach for this setting.

Here, we are concerned with the problem of setting a dynamic threshold above which the value of a variable should be considered as extreme. We are motivated by a well-known

Address for correspondence: Miguel de Carvalho, Department of Statistics, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, Chile.
E-mail: mdecarvalho@mat.puc.cl



(a) (b)

Figure 1. (a) Weekly number of unemployment insurance claims in the US (initial claims). The 2239 weekly observations are seasonally adjusted and range from 7 January 1967 to 28 November 2009; (b) US monthly unemployment rate. The 515 monthly observations are seasonally adjusted and range from January 1967 to November 2009.

economic time series, the weekly number of unemployment insurance claims in the US (henceforth initial claims), often considered as a reference leading indicator able to forestall recessions (Montgomery *et al.*, 1998; Choi and Varian, 2009, and references therein). As it can be observed in Figure 1, there is a natural propensity for the number of initial claims to second-guess the unemployment rate, and this is particularly clear at the end of the observation period where the initial claims peaked before the unemployment rate. Hence, a peaks over threshold analysis could be of interest for assessing the risk of entering into an unemployment surge, given the most recent information available on initial claims. Unemployment is known to behave asymmetrically, in the sense that the probability of a decrease in unemployment, given two previous decreases, is greater than the probability of an increase given two preceding increases (Milas and Rothman, 2008). Unemployment is also supposed to move countercyclically—upward in slowdowns and contractions, and downward in speedups and expansions (Rothman, 1998; Caporale and Gil-Alana, 2008). Thus, the definition of a suitable dynamic threshold could be extremely helpful for recognizing the eruption of those surges and ultimately to help counteract them. As the harshness of some recent unemployment episodes testifies, the understanding of the law of motion of such thresholds is of real value for policy-making.

Classical peaks over threshold methods fail to provide a good modelling solution for the initial claims, since this series is clearly nonstationary. A seminal paper in nonstationary extremes is Davison and Smith (1990), but the quest for alternative modelling approaches is far from complete (Davison and Ramesh, 2000; Hall and Tajvidi, 2000; Chavez-Demoulin and Davison, 2005; Yee and Stephenson, 2007; Padoan and Wand, 2008; Laurini and Pauli, 2009; Eastoe and Tawn, 2009; Northrop and Jonathan, 2011). Most models introduce covariates in the parameters of the threshold model to overcome the lack of stationarity, but some difficulties arise when attempting to apply these to the initial claims data. First, the abovementioned leading attributes of this series make covariate-based approaches non-trivial, as it is difficult to obtain covariates which are also released on a weekly basis, given that most economic data are available monthly or quarterly and released with a large lag. Second, there is a serious risk of establishing spurious associations between the exceedances and the corresponding covariates. As it is well known, the similitude of trending mechanisms in the data can easily lead to spurious regressions, a problem which dates back to Yule (Phillips, 1998; Choi *et al.*, 2008). Third, as recently pointed out by Eastoe and Tawn (2009), covariate-based approaches following Davison and Smith (1990) are unable to preserve one of the most important features of the GPD distribution, viz.: threshold stability, so that the choice of the threshold affects covariate selection; a point process approach (Coles, 2001, §7.3) avoids problems with threshold stability, because its parameters are not threshold-dependent, but it would still require covariates.

We propose to model the initial claims by developing a dynamic threshold modelling approach which avoids the above-mentioned difficulties, and which can be applied to integrated processes of order α, with α denoting any real number; these include fractionally integrated processes which have their roots in the works of Granger and Joyeux (1980) and Hosking (1981). If the process is integrated of order α, then although the time series of interest may be nonstationary, it can be converted into a stationary series by differencing α-times. Since after differencing α-times we obtain stationarity, classical peaks over threshold models can be applied to the resulting series, and binomial series expansions then allow us to build a dynamic threshold for the time series of interest. Given that our approach is linked to the Box–Jenkins method (Box *et al.*, 2008) and the GPD model, we call our strategy the Box–Jenkins–Pareto approach. It is important to stress that our main aim here is not to make inference about the extremes, but rather on setting a threshold to detect movements in a variable. Hence, here we will use extreme value modelling only to choose a level for the threshold of interest, but in another contexts one might be interested in actually using the results of the extreme value analysis to make inferences about levels yet to be observed.

In the next section we provide a preliminary analysis of the data. In §3 we survey the most frequently applied peaks over threshold approaches for stationary and nonstationary time series, and in §4 we introduce our modelling strategy and provide guidelines for implementation. In §5 we examine the weekly number of unemployment insurance claims in the US and explore the connection between its threshold exceedances and the periods of contraction in the US economy.

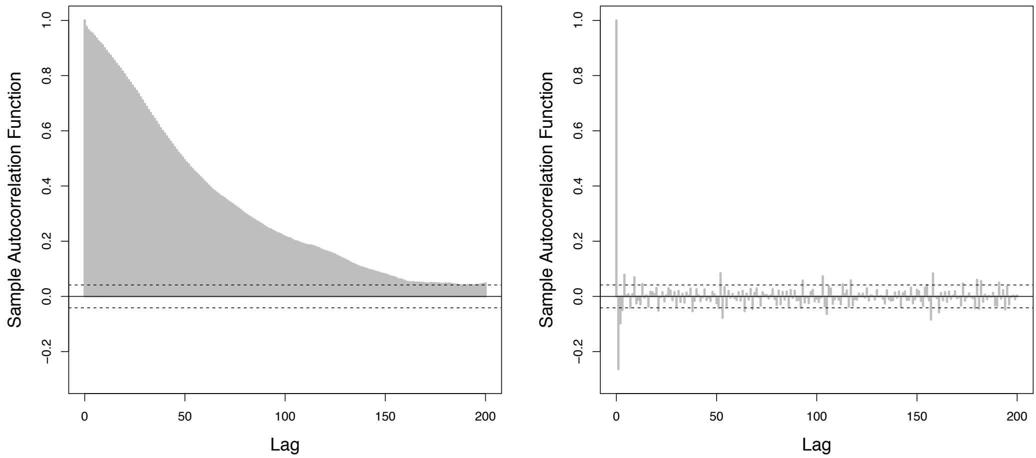
2. Preliminary analysis of initial claims data

To understand the dynamics of the initial claims data $\{Y_1, \ldots, Y_n\}$ in the time domain, we use the correlogram which we define as the set of points $\{(\ell, R_\ell) : \ell \in \{0, \ldots, \ell_{\max}\}\}$, where $\ell_{\max} < n$ is a maximum lag, and R_ℓ is the sample autocorrelation function

$$R_\ell = \frac{\sum_{t=\ell+1}^{n} (Y_t - \overline{Y})(Y_{t-\ell} - \overline{Y})}{\sum_{t=1}^{n} (Y_t - \overline{Y})^2}.$$

The result is plotted in Figure 2 where the dashed lines are based on the pivot $R_\ell \sqrt{n} \rightsquigarrow \mathrm{N}(0,1)$, where '$\rightsquigarrow$' is used to denote weak convergence, which holds under some regularity conditions for realizations of a white noise process (Brockwell and Davis, 2002, Thm 7.2.2). We also plot in Figure 2 the correlogram of the data in first differences. The exploratory analyses reported in Figure 2 clearly suggest long-range dependence in the initial claims, with the correlogram of the data in first differences suggesting a unit root process. In Supporting Information we provide related analyses for the unemployment rate.

In practice we are often faced with the question whether economic time series are stationary (de Carvalho and Júlio, 2012). One of the most used tests for such purpose is the KPSS test (Kwiatkowski *et al.*, 1992) which is based on decomposing the series into a random walk



(a) (b)

Figure 2. Correlogram for (a) the weekly number of unemployment insurance claims in the US (initial claims), for the original data and (b) the data in first differences.

and a white noise series

$$Y_t = H_t + \varepsilon_t, \quad H_t = H_{t-1} + U_t, \quad U_t \overset{\text{iid}}{\sim} (0, \sigma_U^2), \quad \varepsilon_t \overset{\text{iid}}{\sim} (0, \sigma_\varepsilon^2).$$

The interest is in testing the null hypothesis of stationarity, i.e., $\sigma_U^2 = 0$. The KPSS test statistic, at truncation lag parameter ℓ, is given by

$$\text{KPSS}_\ell = \sum_{t=1}^{n} \frac{n^{-2} S_t^2}{\widehat{\sigma}_\ell^2}, \quad S_t = \sum_{i=1}^{t} e_i, \quad t = 1, \dots, n.$$

Here $e_1, \dots, e_n$ are the residuals from the regression $Y_t = \alpha + \varepsilon_t$, and $\widehat{\sigma}_\ell^2$ is a consistent estimator of the long-run variance $\sigma^2 = \lim_{n\to\infty} n^{-1}\text{E}(S_n^2)$, based on the residuals truncated at lag ℓ. We use the so-called Newey–West estimator to estimate σ^2, and take $\ell_{\text{NW}} = 36$, which is the truncation lag parameter chosen according to the Newey–West criterion (Newey and West, 1994). For the initial claims data we obtain $\text{KPSS}_{36} = 0.69$, which yields a p-value of 0.016, so that we reject the null hypothesis of stationarity.

3. The Box–Jenkins–Pareto approach

3.1. Models for stationary time series

Suppose that the time series of interest $\{Y_t\}$ is stationary with univariate marginal survivor function S_Y. Threshold models consider as extreme the observations which exceed a fixed threshold u, and these observations are known as exceedances; to distinguish exceedances from non-exceedances, we use the notation $e_{u,t} = I(Y_t < u)$. The centerpiece of threshold models is based on asymptotic developments (Balkema and de Haan, 1974; Pickands, 1975), which establish that, for a fixed large threshold u, the conditional survivor function of an exceedance by the amount $y > 0$, follows a $\text{GPD}(\varphi_u, \gamma)$, i.e.

$$\text{pr}(Y > u + y \mid Y > u) = \left(1 + \frac{\gamma y}{\varphi_u}\right)_+^{-1/\gamma}, \quad y > 0. \tag{1}$$

Here $\varphi_u > 0$ and $\gamma \in \mathbb{R}$ respectively denote the scale and shape parameters, and $a_+ = \max(a, 0)$. For $\gamma = 0$, (1) should be interpreted by taking the limit $\gamma \to 0$, giving an exponential distribution with parameter $1/\varphi_u$, viz.:

$$\mathrm{pr}(Y > u + y \mid Y > u) = \exp\left(-y/\varphi_u\right), \quad y > 0.$$

After threshold selection has been executed, parameter estimation needs to be conducted. We focus on parameter estimation via likelihood methods; hence, let $\{Y_1, \ldots, Y_n\}$ denote a sample from S_Y, so that the likelihood of the model can be written as

$$\mathscr{L}(S_Y, \varphi_u, \gamma) = \prod_{t=1}^{n} \{1 - S_Y(u)\}^{e_{u,t}} \left[\frac{S_Y(u)}{\varphi_u} \left(1 + \frac{\gamma y_t}{\varphi_u}\right)_+^{-1/\gamma - 1} \right]^{1 - e_{u,t}}. \tag{2}$$

3.2. The Box–Jenkins–Pareto method for nonstationary time series

Data preparation techniques can be quite convenient for subsequent data analysis. One of the most common data preparation methods is given by differencing, i.e., to consider the differences between consecutive observations. The classical Box–Jenkins method is representative of the advantages that differencing can bring into the analysis (Box *et al.*, 2008). Suppose that the nonstationary time series $\{Y_t\}$ can be converted into a stationary series by differencing once, i.e.,

$$(1 - \mathbb{L})Y_t = Z_t, \tag{3}$$

for some stationary series $\{Z_t\}$ with survivor function S_Z. Here $\mathbb{L}$ is the lag operator and $\Delta \equiv (1 - \mathbb{L})$ is the difference operator; a series which satisfies (3) is said to be integrated of order 1, here denoted by I(1). In some cases there is no need to apply the difference operator to preprocess the data, since the process under analysis is already stationary, and a particular case is when we have an I(0) process; formally, a stationary processes with autocovariance function $\gamma_\ell = \mathrm{cov}(Y_t, Y_\ell)$ is an I(0) process, if

$$\sum_{\ell=-\infty}^{\infty} |\gamma_\ell| < \infty.$$

More generally, the series $\{Y_t\}$ is said to be integrated of order α, denoted by I(α), if

$$\Delta^\alpha Y_t = Z_t, \quad \alpha \in \mathbb{R},$$

for some I(0) series $\{Z_t\}$ with survivor function S_Z. A comprehensive discussion on these series can be found in Robinson and Marinucci (2001), and this general class encompasses fractionally integrated processes (Granger and Joyeux, 1980; Hosking, 1981). The long memory parameter α contains information on the stationarity of the sequence: if $\alpha \in [0; 0.5)$ then the series is stationary and mean-reverting; for $\alpha \in [0.5; 1)$ the series is no longer stationary although it is still mean-reverting; finally, if $\alpha \geq 1$ the series is neither stationary nor mean-reverting. A value of $\alpha < 0$ is indicative of mean reversion, but with a type of long memory called 'anti-persistence', where negative autocorrelations die out slowly.

The following functional central limit theorem establishes the link between integrated series and fractional Brownian motion—a continuous stochastic process with known applications in extreme value modelling (Mikosch *et al.*, 2002; Buchamann and Klüppelberg, 2005).

THEOREM 1. *(Sowell, 1990) Let $\{Y_t\}$ be an integrated series of order α, where $-1/2 < \alpha < 1/2$. Suppose that $Z_t \equiv \Delta^\alpha Y_t$ are independent and identically distributed with* $\mathrm{E}(Z_t) = 0$ *and* $\mathrm{E}(|Z_t|^r) < \infty$*, for $r \geq \max\{4, -8\alpha/(1+2\alpha)\}$. Then*

$$\sigma_n^{-1} \sum_{\tau=1}^{\lfloor nt \rfloor} Y_\tau \rightsquigarrow W_\alpha(t), \quad W_\alpha(t) = \frac{1}{\Gamma(\alpha+1)} \int_0^t (t-x)^\alpha \mathrm{d}W(x),$$

where $\sigma_n = \mathrm{var}(\sum_{\tau=1}^{n} Y_\tau)$.

From the extreme value modelling standpoint, the question of interest is the following: suppose that the series of interest $\{Y_t\}$ is nonstationary, but it is $\mathrm{I}(\alpha)$ for some real number α; is it still possible to build directly a threshold model for $\{Y_t\}$? To give an answer to this question, assume by now that the differencing parameter α is a positive integer; later we let α be any real number. Since $\{Y_t\}$ is $\mathrm{I}(\alpha)$, the exceedances of Z, above a fixed large threshold u, can be modelled through a $\mathrm{GPD}(\varphi_u, \gamma)$, i.e.

$$\mathrm{pr}(Z > u + y \mid Z > u) = \left(1 + \frac{\gamma y}{\varphi_u}\right)^{-1/\gamma}, \quad y > 0. \tag{4}$$

Hence, the likelihood of the model is essentially the same as in (2), and for every period t we have that

$$\begin{aligned}
\mathrm{pr}(Z_t > u + y \mid Z_t > u) &= \mathrm{pr}(\Delta^\alpha Y_t > u + y \mid \Delta^\alpha Y_t > u) \\
&= \mathrm{pr}\left[\left\{\sum_{i=0}^{\alpha} \tbinom{\alpha}{i}(-\mathbb{L})^i\right\} Y_t > u + y \,\middle|\, \left\{\sum_{i=0}^{\alpha} \tbinom{\alpha}{i}(-\mathbb{L})^i\right\} Y_t > u\right] \\
&= \mathrm{pr}(Y_t > u_t + y \mid Y_t > u_t).
\end{aligned} \tag{5}$$

Here u_t defines the dynamic threshold given by

$$u_t = u + \sum_{i=1}^{\alpha} \tbinom{\alpha}{i} \{Y_{t-i} I(i \text{ odd}) - Y_{t-i} I(i \text{ even})\}, \tag{6}$$

where $\binom{\alpha}{i} = \Gamma(\alpha+1)/\{\Gamma(i+1)\Gamma(\alpha+i-1)\}$ is the binomial coefficient, and Γ is the gamma function with the conventions $\Gamma(0) = \infty$ and $\Gamma(0)/\Gamma(0) = 1$. The dynamic threshold in (6), is composed of a building block (u) and a time-varying part which uses the previous α values of the time series. From a practical standpoint, this implies that for a sample $\{Y_1, \ldots, Y_n\}$ it is only possible to start the dynamic threshold at the $(\alpha+1)$-observation. This is not however as critical as it might appear since α is finite and independent of n, so that $(\alpha+1)/n = o(1)$, when $n \to \infty$. In the simplest case where the series is difference-stationary with $\alpha = 1$, it holds that $u_t = u + Y_{t-1}$. Hence, the relationship (5) suggests a natural way to construct a dynamic threshold for $\mathrm{I}(1)$ series: obtain u from the first differences of the series of interest, and then add u to the lagged series.

If α is any real number the more general series expansion should be taken into account

$$\Delta^\alpha = \sum_{i=0}^{\infty} \frac{\langle\alpha\rangle_i}{\Gamma(i+1)}(-\mathbb{L})^i, \tag{7}$$

where $\langle\alpha\rangle_i = \alpha(\alpha-1)\cdots(\alpha-i+1)$ is the Pochhammer symbol for the falling factorial. Thus, similarly to (5) we still have

$$\mathrm{pr}(Z_t > u + y \mid Z_t > u) = \mathrm{pr}(Y_t > u_t + y \mid Y_t > u_t),$$

but now the dynamic threshold is

$$u_t = u + \sum_{i=1}^{\infty} \frac{\langle\alpha\rangle_i}{\Gamma(i+1)} \{Y_{t-i} I(i \text{ odd}) - Y_{t-i} I(i \text{ even})\}. \tag{8}$$

In theory, pre-sample shocks could be included in the lag structure—if we had an infinite amount of data—, but in practice we need to suppress them from the lag structure. These different approaches of modelling pre-sample shocks lead to type I and type II fractionally integrated processes, which have been compared in detail by Marinucci and Robinson (1999). For any positive integer α, (7) is tantamount to the classical binomial expansion, so that in this case we recover the threshold given in (6). The more general version of the dynamic threshold in (8) is similar to the one obtained in (6), being also composed by a building block and a time-varying part, but it uses all previous observations, and not just α as before. From a practical standpoint, this implies that for a sample $\{Y_1, \ldots, Y_n\}$, u_t depends on all $t-1$ observations. If $\alpha = 0$ then $\{Y_t\}$ is stationary so that we would expect the peaks over threshold model for stationary time series to hold. From the inspection of (6) we can confirm that this is the case, since $u_t = u$. For completeness we discuss below how the dynamic threshold in (8) can be used for return level modelling; we also discuss the case for integrated series with a polynomial trend.

3.3. The Box–Jenkins–Pareto approach for time series with a polynomial trend

In applications we are often faced with the need to additionally model deterministic trends. Formally, a process $\{Y_t\}$ is said to be integrated of order α, with a polynomial time trend of degree β, if

$$\Delta^{\alpha} Y_t - \lambda t^{\beta} = Z_t, \quad \alpha, \beta \in \mathbb{R},$$

for some stationary I(0) series $\{Z_t\}$ with survivor function S_Z. In this case it holds that

$$\text{pr}(Z_t > u + y \mid Z_t > u) = \text{pr}(Y_t > \upsilon_t + y \mid Y_t > \upsilon_t),$$

where the dynamic threshold is now given by

$$\upsilon_t = u + \lambda t^{\beta} + \sum_{i=1}^{\infty} \frac{\langle \alpha \rangle_i}{\Gamma(i+1)} \{Y_{t-i} I(i \text{ odd}) - Y_{t-i} I(i \text{ even})\}. \tag{9}$$

Hence, the polynomial trend enters additively into the dynamic threshold, and as expected this has now two time-varying components: one due to the trend, and the other due to the memory of the time series. The case where $\alpha = 0$ is again instructive, since then the process is trend-stationary, and hence we obtain $\upsilon_t = u + \lambda t^{\beta}$.

4. The initial claims and the US business cycle

4.1. Data description and statistical packages

We intend to examine what connection the threshold exceedances of the initial claims may have with the contraction periods of the US economy, as dated by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). The period under analysis is from 7 January 1967 to 28 November 2009, and we took the 2239 observations from this seasonally adjusted series from the United States Department of Labor—Employment & Training Administration; the data were downloaded from

`http://www.ows.doleta.gov`

The US Department of Labor revise the data periodically, and the version we use here was downloaded on December 2010. One could think of using the exceedances resulting from the dynamic threshold scheme introduced above, as an indicator of whether an economy is entering or crossing a recession period. Several reasons anticipate however the difficulties with such inquiry, and one of such complications lies in the data itself. As pointed by the Business Cycle Dating Committee of the NBER (see Frequently Asked Questions NBER, 2008), there is marked week-to-week noise in the initial claims; cf. Figure 1.

All the analyses related with extreme value modelling were conducted using the R package `evd} \citep{Stephenson:02,Stephenson:06}.}`

4.2. Box–Jenkins–Pareto analysis

To apply the Box–Jenkins–Pareto approach we first need to estimate the long memory parameter α. There has long been an interest in fractionally integrated models and in the estimation of the fractional differencing parameter; for a recent review see Gil-Alana and Hualde (2009). Among the estimation procedures available, parametric and semi-parametric approaches are the most employed in practice. In the former, a full parametric model is specified, and so there is the risk of misspecification, which can yield biased estimates of α (Fox and Taqqu, 1986; Sowell, 1992). A semi-parametric approach to estimate the fractional differencing parameter is here pursued. We use the well-known GPH estimator (Geweke and Porter-Hudak, 1983) which corresponds to the least squares estimate of slope, in the log-periodogram regression

$$\log\{\mathbb{I}(\omega_j)\} = a + \alpha \log\{4\sin^2(\omega_j/2)\}^{-1} + \varepsilon_j, \quad \varepsilon_j \sim (0, \pi^2/6), \quad j = 1, \ldots, \lfloor m \rfloor, \tag{10}$$

where $\omega_j = 2\pi j/n$. A practical problem in its implementation is the selection of an upper bound for the number of frequencies m to be used in the regression—a choice that entails a bias–variance trade-off. Geweke and Porter-Hudak suggested using $m = n^{1/2}$, and this is the rule most often used in practice. Applying this rule to the initial claims data yields $\lfloor m \rfloor = 47$ and $\widehat{\alpha} = 0.96$, with a standard error of 0.11. Since the choice of the number of frequencies to be used in the regression is not clear cut (Hurvich and Deo, 1999), we also present the GPH estimate for a wide range of frequencies; following Perron and Qu (2010), we consider m ranging from 10 up to $n^{3/4}$. From Figure 3, except for the case where only extremely low frequencies are used—where the GPH estimate reveals to be unstable—, the GPH estimate is always close to one.

As usual, threshold selection is a debatable step. If a too low threshold is selected then the asymptotic rationale of the model is not justified and bias is generated. On the other hand, if a too high threshold is chosen few exceedances are available so that higher variance is obtained; detailed recommendations on threshold selection can be found in Bermudez *et al.* (2001). We use the mean residual life plot and plotted parameter estimates of the peaks over threshold model, of the preprocessed data $Z_t = \Delta^{0.96} Y_t$, at a variety of thresholds. As it can be observed in Figure 4 the estimate of the tail index estimate is stable if small perturbations are induced in the fixed threshold of $u^+ = 48$, and a similar analysis for the left tail suggests a fixed threshold of $u^- = -38$. The analysis was supplemented by quantile plots, which are given in Figure 5 along with their simulation envelopes, and which provide evidence supporting a reasonably good fit of the model.

Extra diagnostics and some numerical experiments with simulated data can be found in Supporting Information. We also discuss in Supporting Information, conditions under which it is possible to obtain asymptotic normality for the time-varying threshold, at each period t. These results are however of limited applied interest, and in practice we prefer to use a bootstrap scheme based on Arteche and Orbe (2009), so that for each period t, we:

I. estimate $\widehat{a}$ and $\widehat{\alpha}$ by least squares in (10), and obtain the residuals $e_j = \widehat{a} - \widehat{\alpha}\log\{4\sin^2(\omega_j/2)\}^{-1}$, for $j = 1, \ldots, \lfloor m \rfloor$;

II. obtain bootstrap samples $e^{\star}_{bj}$ from the the empirical distribution function of e_j, and the corresponding bootstrap dependent variable $\mathbb{I}^{\star}(\omega_j)$ for $j = 1, \ldots, \lfloor m \rfloor$ and $b = 1, \ldots, B$;

III. obtain $\alpha^{\star}_b$ from the bootstrap regressions, for $b = 1, \ldots, B$; obtain the bootstrap distribution of $\widehat{u}^{\star}_{tb} \mid Y_{t-1}, \ldots, Y_1$ from

$$u + \sum_{i=1}^{t-1} \frac{\langle \widehat{\alpha}^{\star}_b \rangle_i}{\Gamma(i+1)} \left\{ Y_{t-i} I(i \text{ odd}) - Y_{t-i} I(i \text{ even}) \right\}.$$

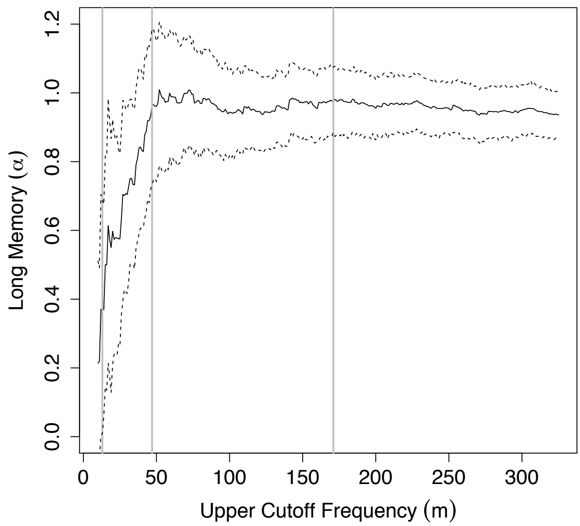


Figure 3. The solid line represents the GPH estimates of α with m ranging from 10 up to $n^{3/4}$; the gray vertical lines correspond to $n^{1/3}$, $n^{1/2}$, and $n^{2/3}$, and the dashed lines represent 95% pointwise confidence bands, which were obtained by the bootstrap.

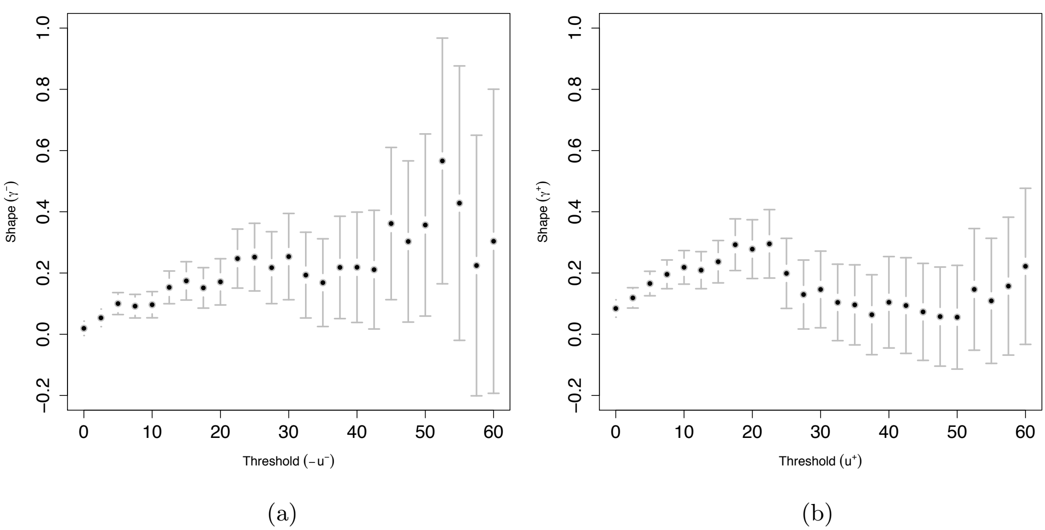
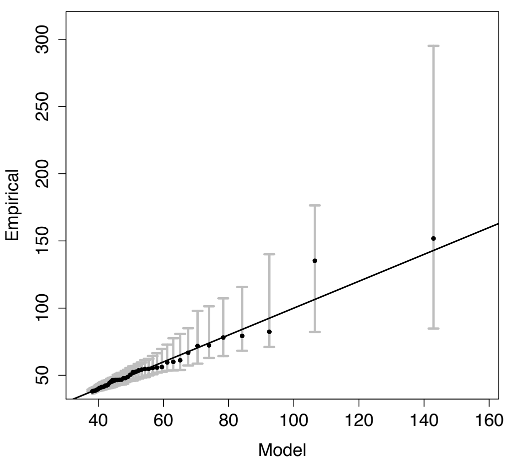


Figure 4. Maximum likelihood estimates of the shape parameters and corresponding 95% confidence intervals over a grid of thresholds; left and right tails are in (a) and (b), respectively.

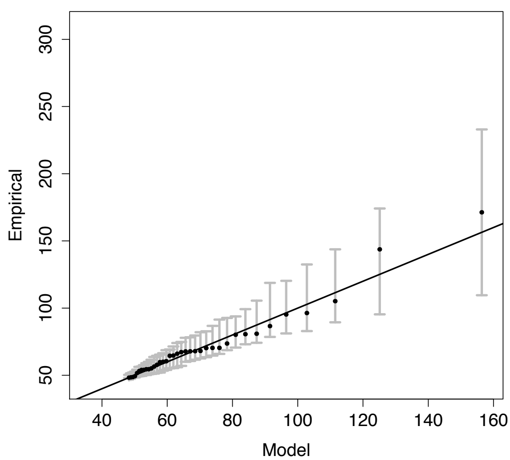
Table 1. Maximum likelihood estimates for the peaks over threshold analysis for left ($-$) and right ($+$) tails, considering the thresholds $u^- = -38$ and $u^+ = 48$; standard errors are in parentheses

	Left tail		Right tail	
α	$\varphi^-_{u^-}$	γ^-	$\varphi^+_{u^+}$	γ^+
0.96	16.09	0.04	21.45	0.07
(0.11)	(3.22)	(0.13)	(4.99)	(0.16)
1	12.87	0.24	21.14	0.08
(—)	(2.89)	(0.17)	(4.96)	(0.17)

The bootstrap approach of Arteche and Orbe (2009) was also used to obtain the pointwise confidence bands for the GPH estimates in Figure 3. In Figure 6 we plot the time-varying threshold for $\widehat{\alpha} = 0.96$ and its corresponding 95% pointwise confidence bands which were constructed using the bootstrap, with $B = 1000$; we present the results for a short window, here given by the first ten weeks of data, and there it can be observed that a threshold exceedance occurs on week 9. In Table 1 we report the maximum likelihood estimates for peaks over threshold analysis for left and right tails, for the cases where the long memory parameter is 0.96 and 1. The estimates of the shape parameters are larger for $\alpha = 1$, whereas the estimates of the scale parameter are larger for $\alpha = \widehat{\alpha} = 0.96$. The standard errors in the shape and scale parameters are however too large to assess if these ranks are actually significant.



(a)



(b)

Figure 5. Quantile plots for the fits of the peaks over threshold model and corresponding simulated 95% confidence intervals corresponding to: (a) left tail; (b) right tail.

4.3. Economic interpretation and mirror filtered exceedances

To give some interpretation to the sequence of exceedances generated by the dynamic threshold, we introduce in the subsequent figures shaded areas representing the US economic activity contractions dated by the Business Cycle Dating Committee of the NBER. It should

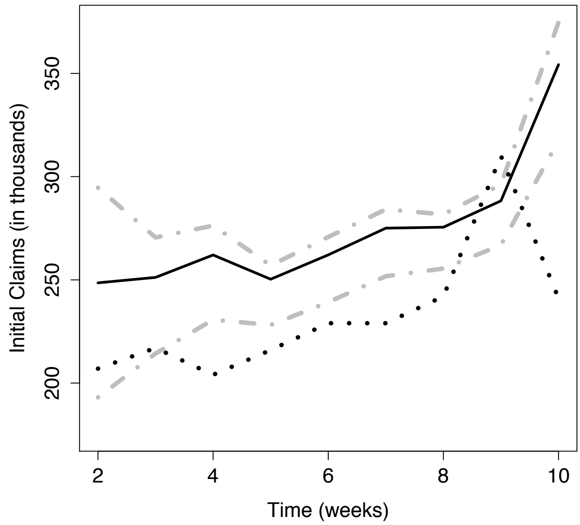
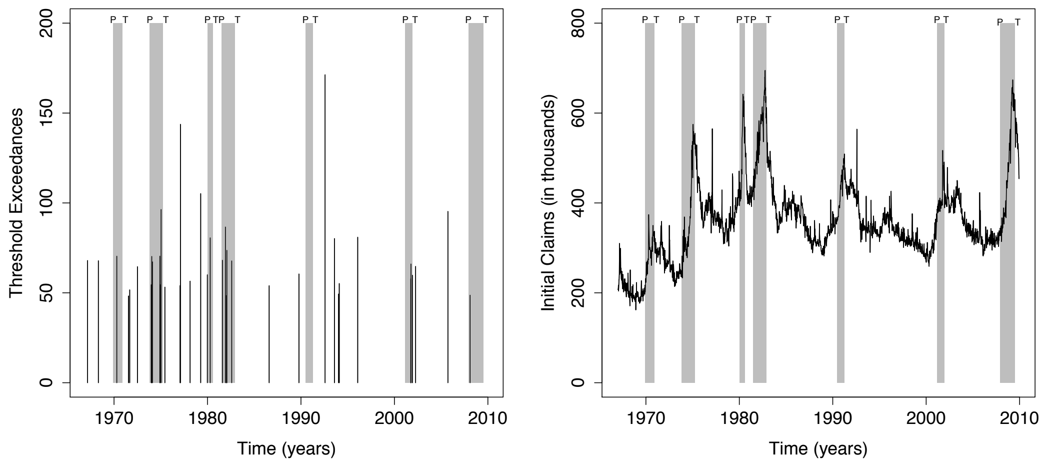


Figure 6. The solid line represents the dynamic threshold constructed according to (8); the 95% pointwise confidence bands constructed by the bootstrap, and the initial claims are respectively represented by the dashed–dotted and dashed lines.

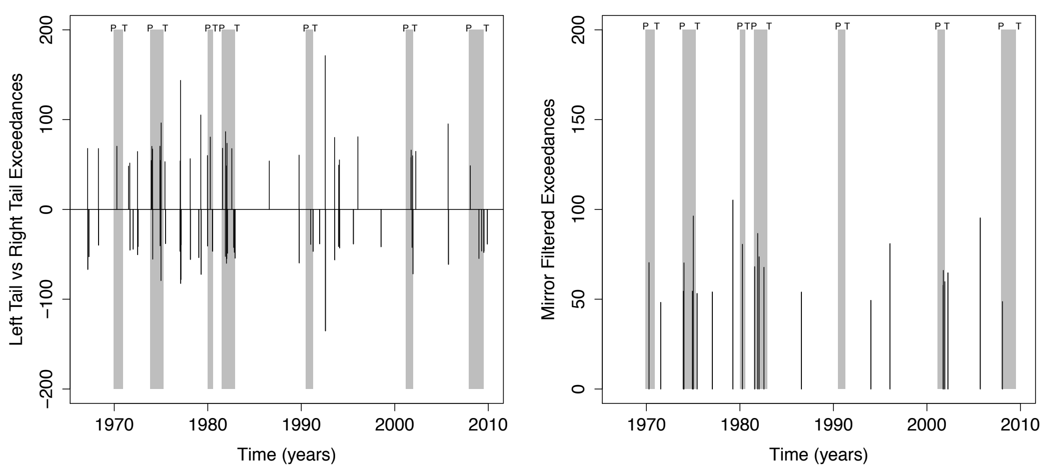
be stressed that our aim here is not to design an optimal alarm mechanism—in the sense of Antunes *et al.* (2003)—, but merely to provide some economic interpretation to the threshold exceedances found by using our dynamic threshold procedure. We also note that our analysis is ex-post, so that there is really no forecasting involved. Seven peak (P) to trough (T) movements occurred from January 1967 to November 2009. Thus, during the period under analysis seven contractions were acknowledged by the NBER Business Cycle Dating Committee, viz.: *i*) December 1969–November 1970; *ii*) November 1973–March 1975; *iii*) January 1980–July 1980; *iv*) July 1981–November 1982; *v*) July 1990–March 1991; *vi*) March 2001–November 2001; *vii*) December 2007–June 2009. Observe further that there is some lag in the identification of peaks by NBER. For example, the economic activity peak of December 2007 was only determined in December 2008 (NBER, 2008).

Figure 7 represents the threshold exceedances and the original series. We now assess the information content that the threshold exceedances of the initial claims possess for tracking contraction periods. From the inspection of Figure 7 we can ascertain that among the 2239 weekly observations such mechanism would have been activated only 37 times. It is promising that such naive mechanism is consistent with several contraction episodes and particularly with the eruption of the latest economic activity peak determined by NBER. This is reinforced by the fact that in only 17.6% of the periods under analysis contractions occurred, so that it is substantially more difficult to spot recessive periods simply by chance. The analysis of Figure 7 also reveals however that several exceedances occurred during expansions. As argued above, it is recognized by the NBER (See Frequently Asked Questions NBER, 2008) that there is a noticeable week-to-week noise in the initial claims series which complicates its analysis. As it can be observed in Figure 7 the larger exceedances in (a) correspond to isolated spikes in (b) so that they are most probably due to week-to-week noise. In general, these spikes are immediately reverted in the following week. Therefore, one possible way to sieve plausible exceedances from noisy ones is to inspect which exceedances were followed in the next week by a left tail exceedance.



(a) (b)

Figure 7. (a) Threshold exceedances; (b) Weekly number of unemployment insurance claims in the US (initial claims). Shaded areas represent the US economic activity contractions dated by the Business Cycle Dating Committee of the NBER.



(a) (b)

Figure 8. (a) Mirror plot; (b) Mirror filtered exceedances.

Figure 8 (a) depicts the right and left tail exceedances—a representation which we denote as the mirror plot. The analogy here is that the lines corresponding to noisy exceedances should be immediately followed by left tail exceedances creating the visual effect of a mirror image. The mirror plot can then be thought as an exploratory tool for examining which right tail exceedances are followed by left tail exceedances in the next week. Observe that the filtering procedure suggested by the mirror plot is congruous with the earlier discussed dynamic asymmetry, according to which unemployment exhibits abrupt increases in opposition to longer and gradual declines (Milas and Rothman, 2008). In particular this implies that right tail exceedances are not expected to be immediately followed by left tail exceedances. The right tail exceedances which are not followed by a left tail exceedance in the upcoming week are represented in Figure 8 (b) and are here denoted as mirror filtered exceedances. Formally, we define mirror filtered exceedances as the sequence of point masses

$$Z_t^{\mathrm{F}} = Z_t \delta_t(T), \quad t \in T,$$

where $T = \{\tau : Z_\tau > u^+, Z_{\tau+1} < u^-\}$ and $\delta_t(\cdot)$ denotes the Dirac measure at t. It may appear from observing the mirror plot that some other exceedances should also disappear, but as it can be observed in Table 1 in Supporting Information these exceedances are actually reflected after more than one week. To additionally filter these exceedances we can extend the set T, so that more lags are taken into account. The number of mirror filtered exceedances is thus $|T| = 23$, from which 14 occurred during contraction periods and 9 during expansion periods. It is important to note that in only circa 4/23 of the periods under analysis contractions occurred. This implies that it is much more difficult to randomly spot contraction periods, so that a proportion of 14/23 is considerably satisfying; this is reinforced by a one-tailed binomial test against the alternative H_1 : proportion $>$ 4/23, which yields a p-value of 3×10^{-6}. It should also be pointed out that two of the mirror filtered exceedances which occurred out of contraction periods are only a few weeks apart from the trough, and among the remainder only five are clearly distant from any contraction period.

5. Discussion

This paper discusses an approach for setting a dynamic threshold for a leading economic indicator, and explores the connection of its exceedances with the contraction periods of the US economy. Our approach is based on fractionally integrated processes, which allow us to construct a time-varying threshold that links the preprocessed series with the series of interest. It is important to stress that here, we use extreme value modelling only to choose a level for the threshold (the functional form of our dynamic threshold does not rely on the extreme value theory), but in another application, one might be interested in actually using the results of the extreme value analysis to make inferences about levels yet to be observed.

The case of fractionally integrated processes with a polynomial trend requires alternative estimation procedures, and a natural one is given by the extended exact local Whittle estimator of Shimotsu (2010). If one prefers to work on a scale which has a direct interpretation—say first differences—, the value of α can still be indicative of the level of misspecification we incur by working on such scale. The closer the α is to zero, the more reasonable it may be to apply a peaks over threshold model to Y_t directly. Similarly, a value of $\alpha \in (0.5, 1]$ suggests that it is not sensible to work with Y_t, and that the closer the α is to 1, the better the approximation provided by ΔY_t.

Acknowledgments

We thank the editor, an associate editor, and two reviewers for their helpful comments and suggestions on an earlier version of this paper. We are grateful to Anthony Davison, Valérie

Chavez-Demoulin, Vanda Inácio, Paulo Júlio, and Paulo M. M. Rodrigues for helpful discussions and recommendations that led to a significant improvement of this paper. We also thank comments from participants at a 2010 SAMSI workshop chaired by Richard Smith. Part of this paper was written while M. de C. was visiting the Economics Research Department of *Banco de Portugal*, and revised during his stay at the *Ecole Polytechnique Fédérale de Lausanne.* This research is partially supported by national grants through the *Fundação para a Ciência e a Tecnologia* (FCT), Portugal, under the projects PEst-OE/MAT/UI0297/2011, PEst-OE/MAT/UI0006/2011, and PTDC/MAT/118335/2010.

References

Antunes, M., Turkman, M. A. and Turkman, K. F. (2003) A Bayesian approach to event prediction. *J. Time Ser. Anal.*, **24**, 631–646.

Arteche, J. and Orbe, J. (2009) Using the bootstrap for finite sample confidence intervals of the log periodogram regression. *Computnl Statist. Data Anal.* **53**, 1940–1953.

Balkema, A. A. and de Haan, L. (1974) Residual life time at great age. *Ann. Probab.*, **2**, 792–804.

Bermudez, P., Turkman, M. A. and Turkman, K. F. (2001) A predictive approach to tail probability estimation. *Extremes*, **4**, 295–314.

Brockwell, P. J. and Davis, R. A. (2002) *Introduction to Time Series and Forecasting.* 2nd edn. New York: Springer.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. (2008) *Time Series Analysis: Forecasting and Control.* 4th edn. New Jersey: Wiley.

Buchamann, B. and Klüppelberg, C. (2005) Maxima of stochastic processes driven by fractional Brownian motion. *Adv. Appl. Probab.*, **37**, 743–764.

Caporale, G. M. and Gil-Alana, L. (2008) Modelling the US, UK and Japanese unemployment rates: fractional integration and structural breaks. *Computnl Statist. Data Anal.*, **52**, 4998–5013.

Chavez-Demoulin, V. and Davison, A. C. (2005) Generalized additive modelling of sample extremes. *J. R. Statist. Soc.* C, **54**, 207–222.

Choi, C-Y., Hu, L. and Ogaki, M. (2008) Robust estimation for structural spurious regressions and a Hausman-type cointegration test. *J. Econometrics*, **142**, 327–351.

Choi, H. and Varian, H. (2009) Predicting initial claims for unemployment benefits. *Technical Report—Google.*

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values.* London: Springer.

Davis, R. A. and Mikosch, T. (2009) The extremogram: a correlogram for extreme events. *Bernoulli*, **15**, 977–1009.

Davison, A. C. and Ramesh, N. I. (2000) Local likelihood smoothing of sample extremes. *J. R. Statist. Soc.* B, **62**,191–208.

Davison, A. C. and Smith, R. (1990) Models for exceedances over high thresholds (with discussion). *J. R. Statist. Soc.* B, **52**, 393–442.

de Carvalho, M. and Júlio, P. (2012) Digging out the PPP hypothesis: an integrated empirical coverage. *Empir. Econ.*, **42**, 713–744.

Eastoe, E. F. and Tawn, J. A. (2009) Modelling non-stationary extremes with application to surface level ozone. *J. R. Statist. Soc.* C, **58**, 25–45.

Einmahl, J. H. J. and Magnus, J. R. (2008) Records in athletics through extreme-value theory. *J. Am. Statist. Assoc.*, **103**, 1382–1391.

Fox, R. and Taqqu, M. S. (1986) Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, **14**, 517–532.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long-memory time series models. *J. Time Ser. Anal.*, **4**, 221–238.

Gil-Alana, L. and Hualde, J. (2009) Fractional integration and cointegration. An overview and an empirical application. *Palgrave Handbook in Econometrics*, **2**, (Eds. C. Mills and K. Patterson) pp. 434–472, MacMillan Publishers.

Granger, C. W. J. and Joyeux, R. (1980) An introduction to long memory time series and fractional differencing. *J. Time Ser. Anal.*, **1**, 5–39.

Hall, P. and Tajvidi, N. (2000) Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.*, **15**, 153–167.

Hosking, J. R. M. (1981) Fractional differencing. *Biometrika*, **68**, 165–176.

Hurvich, C. M. and Deo, R. S. (1999) Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long memory time series. *J. Time Ser. Anal.*, **20**, 331–341.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) How sure are we that economic time series have a unit root? *J. Econometrics*, **54**, 159–178.

Laurini, F. and Pauli, F. (2009) Smoothing sample extremes: the mixed model approach. *Computnl Statist. Data Anal.*, **53**, 3842–3854.

Marinucci, D. and Robinson, P. M. (1999) Alternative forms of fractional Brownian motion. *J. Statist. Plann. Inf.*, **80**, 111–122.

Mikosch, T., Resnick, S., Rootzén, H. and Stegeman, A. W. (2002) Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, **12**, 23–68.

Milas, C. and Rothman, P. (2008) Out-of-sample forecasting of unemployment rates with pooled STVECM forecasts. *Int. J. Forecasting*, **24**, 101–121.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S. and Tiao, G. C. (1998) Forecasting the US unemployment rate. *J. Am. Statist. Assoc.*, **93**, 478–493.

NBER (2008) Determination of the December peak in economic activity. Announcement of the NBER's Business Cycle Dating Committee (`http://www.nber.org/cycles/`).

Newey, W. K. and West, K. D. (1994) Automatic lag selection in covariance matrix estimation. *Rev. Econ. Studies*, **61**, 631–653.

Northrop, P. J. and Jonathan, P. (2011) Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights (with discussion). *Environmetrics*, **22**, 799–809.

Padoan, S. A., Ribatet, M. and Sisson, S. A. (2010) Likelihood-based inference for max-stable processes. *J. Am. Statist. Assoc.*, **105**, 263–277.

Padoan, S. A. and M. P. Wand (2008) Mixed model-based additive models for sample extremes. *Statist. Probab. Lett.*, **78**, 2850–2858.

Perron, P. and Qu, Z. (2010) Long-memory and level shifts in the volatility of stock market return indices. *J. Bus. Econ. Statist.*, **28**, 275–290.

Phillips, P. C. B. (1998) New tools for understanding spurious regressions. *Econometrica*, **66**, 1299–1325.

Pickands, J. (1975) Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 119–131.

Poon, S.-H., Rockinger, M. and Tawn, J. A. (2003) Modelling extreme-value dependence in international stock markets. *Statist. Sin.*, **13**, 929–953.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Robinson, P. M. and Marinucci, D. (2001) Narrow-band analysis of nonstationary processes. *Ann. Statist.*, **29**, 947–986.

Rothman, P. (1998) Forecasting asymmetric unemployment rates. *Rev. Econ. Statist.*, **80**, 164–168.

Sowell, F. (1990) The fractional unit root distribution. *Econometrica*, **58**, 495–505.

Sowell, F. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *J. Econometrics*, **53**, 165–188.

Shimotsu, K. (2010) Exact local Whittle estimation of fractional integration with unknown mean and time trend. *Econometric Th.*, **26**, 501–540.

Stephenson, A. (2002) evd: extreme value distributions. *R News*, **2**, 31–32.

Stephenson, A. and Gilleland, E. (2006) Software for the analysis of extreme events: The current state and future directions. *Extremes*, **8**, 87–109.

Turkman, K. F., Turkman, M. A. and Pereira, J. M. (2010) Asymptotic models and inference for extremes of spatio-temporal data extremes. *Extremes*, **13**, 375–397.

Yee, T. W. and A. G. Stephenson (2007) Vector generalized linear and additive extreme value models. *Extremes*, **10**, 1–19.