DOCUMENT DE TREBALL

XREAP2015-01

ESTIMATING EXTREME VALUE CUMULATIVE DISTRIBUTION FUNCTIONS USING BIAS-CORRECTED KERNEL APPROACHES

Catalina Bolancé (Riskcenter-IREA, XREAP)
Zuhair Bahraoui (Riskcenter-IREA, XREAP)
Ramon Alemany (Riskcenter-IREA, XREAP)

XARXA DE REFERÈNCIA EN ECONOMIA APLICADA

Estimating extreme value cumulative distribution functions using bias-corrected kernel approaches

Catalina Bolancé, Zuhair Bahraoui and Ramon Alemany
Department of Econometrics, Riskcenter-IREA, University of Barcelona,
Av. Diagonal, 690, 08034 Barcelona, Spain

January 21, 2015

Abstract

We propose a new kernel estimation of the cumulative distribution function based on transformation and on bias reducing techniques. We derive the optimal bandwidth that minimises the asymptotic integrated mean squared error. The simulation results show that our proposed kernel estimation improves alternative approaches when the variable has an extreme value distribution with heavy tail and the sample size is small.

Keywords: transformed kernel estimation, cumulative distribution function, extreme value distribution.

1 Introduction

Estimating the cumulative distribution function (cdf) is a fundamental goal in many fields in which analysts are interested in estimating the risk of occurrence of a particular event, for example, the probability of a catastrophic accident or the probability of a major economic loss. Similarly, in risk quantification, risk measurements are usually expressed in terms of the cdf, a good example being the distortion risk measures proposed in Wang (1995, 1996).

Specifically, risk quantification concentrates in the highest values of the domain of the distribution, where sample information is scarce and it is, therefore, necessary to extrapolate the behaviour of the cdf, even above the maximum observed. To extrapolate the distribution we can use parametric models or, alternatively, we can use a nonparametric estimation. In this paper, we propose a nonparametric estimator of the cdf that allows us to extrapolate the behaviour of the cdf with greater accuracy than is possible with existing methods.

A naive nonparametric estimator of the cdf is the empirical distribution. It is known that the empirical distribution is an unbiased estimator of cdf. However, the empirical distribution is inefficient when data are scarce. A nonparametric alternative for estimating the cdf is the kernel estimator. This is more efficient than the empirical distribution but it is, nevertheless, a biased estimator. Furthermore, both the empirical distribution and the kernel estimator of the cdf are inefficient when the shape of the distribution is right skewed and it has a longer right tail, i.e., it belongs to a certain family of extreme value distributions (EVD): the Gumbel or Fréchet types. In these cases, although we have a large sample size, the number of observations in the highest values of the domain of the distribution is small. This kind of distribution is very common in microeconomic, financial and actuarial data, where economic quantities are measured, e.g., costs, losses and wages. Likewise, there are other fields such as demography, geology or meteorology, where the observed phenomena are distributed following an EVD (see, for example, Reiss and Thomas, 1997). In this study, we develop a bias-corrected transformed kernel estimator of the cdf that is more accurate than the bias-corrected classical kernel estimator.

With the aim of reducing the bias of the classical kernel estimator (CKE) of the cdf, Kim et al. (2006), based on Choi and Hall (1998), proposed a bias reducing technique, henceforth the bias-corrected classical kernel estimator (BCCKE). Alternatively, Alemany et al. (2013) proved that using the transformed kernel estimator of the cdf, the bias and variance of the CKE could be reduced and they proposed a new estimator based on two transformations, the double transformed kernel estimator (DTKE). However, this estimator has asymptotic properties and needs a large sample size to obtain better results than alternative approaches. In this study, we analyse the properties of the DTKE of the cdf by incorporating the finite sample bias correction proposed by Kim et al. (2006). We refer to this new estimator as the bias-corrected double transformed kernel estimator (BCDTKE).

Estimating the smoothing parameter associated with kernel estimations is also a challenge when the data are generated by an extreme value distribution. When using the two most popular automatic methods. i.e., plug-in and cross-validation, the optimal value frequently degenerates to zero. An alternative for calculating the smoothing parameter is the rule-of-thumb value (Silverman, 1986), which is based on a reference distribution. Using the proposed BCDTKE we can estimate the exact rule-of-thumb value based on a known distribution.

The use of nonparametric methods is based on the lack of information about the theoretical distribution associated with the random variable under analysis. This distribution might match one of those belonging to a subfamily of EVDs: Type I (Gumbel) or Type II (Fréchet). Moreover, the distribution might be a mixture of two or more EVDs. An important goal of this study is to analyse the domain of attraction of different mixtures of EVDs. In section 2 we carry out this analysis. In section 3 we describe the BCCKE of cdf and we propose some new results related to the asymptotically optimal smoothing parameter. These results are then used in section 4, where we describe a new estimator based on transformations and bias correction.

In section 5, we show the results of a simulation study. We conclude in section 6.

2 Maximum domain of attraction of mixtures of extreme value distributions

In this section we prove some results related to the maximum domain of attraction (MDA) of some mixtures of EVDs. The expression of the cdf of a generalised EVD is (see, Jenkinson, 1955):

$$\begin{array}{ll} G_\xi(x,\mu,\sigma) = \exp\left\{-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right\} & \text{if } \xi \neq 0 \\ G_\xi(x,\mu,\sigma) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\} & \text{if } \xi = 0 \end{array} \tag{1}$$

and its mean is:

$$E(X) = \begin{cases} \mu + \sigma\frac{\Gamma(1-\xi)-1}{\xi} & \text{if} \quad \xi \neq 0, \xi < 1 \\ \mu + \sigma\gamma & \text{if} \quad \xi = 0 \\ \infty & \text{if} \quad \xi \geq 1 \end{cases}, \tag{2}$$

where $\Gamma(\cdot)$ is Euler's gamma function and γ is Euler's constant. The MDA of G_ξ ($MDA(G_\xi)$) depends on the shape parameter ξ. In the expression (1) when $\xi = 0$ a Gumbel type EVD is obtained and when $\xi > 0$ the result is a Fréchet type EVD.

We define the right end point of G as $r(G) = \sup(x|G(x) < 1)$. We know that if two distributions F and G are such that $r(G) = r(F)$ then:

$$\lim_{x \uparrow r(F)} \frac{\bar{F}(x)}{\bar{G}(x)} = c,$$

for some constant $0 < c < \infty$, where $\bar{F}(x) = 1 - F(x)$ and $\bar{G}(x) = 1 - G(x)$. In this case F and G have the same MDA, furthermore, F and G are tail equivalent if (see, for example, Embrechts et al., 1997):

$$\lim_{x \uparrow r(F)} \frac{\bar{F}(x)}{\bar{G}(x)} = 1.$$

Theorem 1 *Let F be a cdf that is expressed as $F(x) = \sum_{i=1}^m p_i F_i(x)$, with $\sum_{i=1}^m p_i = 1, \forall p_i > 0$, if every $F_i \in MDA(G_{\xi_i})$, with $\xi_i > 0$ (Fréchet), then $F \in MDA(G_{\xi_M})$, where $\xi_M = max(\xi_1, \ldots, \xi_m)$.*

Proof 1 *We know that if* $F_i \in MDA(G_{\xi_i})$*,* $\forall i = 1, ..., m$*, with* $\xi_i > 0$ *(Fréchet), then* $\bar{F}_i(x) = x^{-\frac{1}{\xi_i}} L_i(x)$*, where* L_i *is a slowly varying function*[1] *and*

$$\begin{aligned}
\bar{F}(x) = 1 - F(x) &= \sum_{i=1}^{m} p_i \bar{F}_i(x) \\
&= \sum_{i=1}^{m} p_i x^{-\frac{1}{\xi_i}} L_i(x) \\
&= x^{-\frac{1}{\xi_M}} \sum_{i=1}^{m} p_i x^{(\frac{1}{\xi_M} - \frac{1}{\xi_i})} L_i(x) \\
&= x^{-\frac{1}{\xi_M}} p_M L_M(x) + x^{-\frac{1}{\xi_M}} \sum_{i \neq M} p_i x^{(\frac{1}{\xi_M} - \frac{1}{\xi_i})} L_i(x) \\
&\sim x^{-\frac{1}{\xi_M}} p_M L_M(x).
\end{aligned}$$

The previous result is obtained observing that $x^{(\frac{1}{\xi_M} - \frac{1}{\xi_i})} L_i(x) \rightarrow_{x \rightarrow \infty} 0$*. We conclude that* F *and* F_M *are tail equivalents.*

Theorem 2 *(Sufficient condition) If* $j \in \{1, \ldots, m\}$ *is such that* $\lim_{x \rightarrow \infty} \frac{\bar{F}_i(x)}{\bar{F}_j(x)} = A < \infty$*, with* $j \neq i$*, then if* $F(x) = \sum_{i=1}^{m} p_i F_i(x) \in MDA(G_\xi)$*, with* $\xi > 0$*, then* $F_j \in MDA(G_\xi)$*.*

Proof 2 *To prove Theorem 2 we start with the definition of regular variation. A positive Lebesgue measurable function* L *on* $(0, \infty)$ *is a regular variation at infinity with index* $\alpha \in \mathbb{R}$ *if:*

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = t^\alpha, \quad t > 0. \tag{3}$$

Then, $F \in MDA(G_\xi)$ *with* $\xi > 0$ *(Fréchet), if* $\bar{F}$ *is a regular variation with index* $-\frac{1}{\xi}$*, namely:*

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = t^{-\frac{1}{\xi}}, \quad t > 0. \tag{4}$$

[1] A positive Lebesgue measurable function L on $(0, \infty)$ is slowly varying if

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad t > 0.$$

We have:

$$\begin{aligned}
\lim_{x\to\infty}\frac{\bar{F}(tx)}{\bar{F}(x)} &= t^{-\xi}\\
\lim_{x\to\infty}\frac{\sum_{i=1}^{m}p_i\bar{F}_i(tx)}{\sum_{i=1}^{m}p_i\bar{F}_i(x)} &= t^{-\xi}\\
\lim_{x\to\infty}\frac{\left[\sum_{i\neq j}^{m}p_i\frac{\bar{F}_i(tx)}{\bar{F}_j(tx)}+p_j\right]\bar{F}_j(tx)}{\left[\sum_{i\neq j}^{m}p_i\frac{\bar{F}_i(x)}{\bar{F}_j(x)}+p_j\right]\bar{F}_j(x)} &= t^{-\xi},
\end{aligned}$$

taking into account the limit in the interior of the brackets and considering the condition $\lim_{x\to\infty}\frac{\bar{F}_i(x)}{\bar{F}_j(x)} = A < \infty$ *we deduce:*

$$\lim_{x\to\infty}\frac{\bar{F}_j(tx)}{\bar{F}_j(x)} = t^{-\xi},$$

then F_j is a Fréchet type EVD.

Theorem 3 *Let F be a cdf that is expressed as $F(x)=\sum_{i=1}^{m}p_iF_i(x)$, with $\sum_{i=1}^{m}p_i=1$, $\forall p_i>0$, if $j\in\{1,\ldots,m\}$ is such that $F_j\in MDA(G_{\xi_j})$, with $\xi_j>0$ (Fréchet), and F_i $\forall i\neq j$ are Lognormal distributions then $F\in MDA(G_{\xi_j})$.*

Proof 3 *Firstly we note that:*

$$\sup(x|F_j(x)<1)\subset\sup(x|\sum_{i=1}^{m}p_iF_i(x)<1)$$

and the right end point of F is $r(F)=\sup(x|F(x)<1)=\infty$. Besides, we have:

$$\bar{F}(x)=\sum_{i\neq j}p_i\bar{F}_i(x)+p_j\bar{F}_j(x),$$

where:

$$\begin{aligned}
\bar{F}_j(x) &= x^{\frac{-1}{\xi_j}}L(x),\ \text{where } L(x) \text{ is slowly varying function,}\\
\bar{F}_i(x) &= \bar{\Phi}(\frac{log(x)-\mu}{\sigma}),\ \text{where } \Phi \text{ is Normal standard distribution,}\\
\bar{F}_i(x) &\sim \frac{\varphi(\frac{log(x)-\mu}{\sigma})}{(\frac{log(x)-\mu}{\sigma})},\ \text{where } \varphi \text{ is Normal standard density,}
\end{aligned} \tag{5}$$

the last term in (5) is deduced applying l'Hôpital's rule to $\frac{x\bar{\Phi}(t)}{\varphi(t)}$*, resulting in* $\bar{\Phi}(t) \sim \frac{\varphi(t)}{t}$ *when* t *is high. If* $\xi_j > 0$ *we can find* $\alpha > 0$ *such that* $\frac{1}{\xi_j} + \alpha > 0$ *and*

$$\begin{aligned}\frac{\bar{F}_i(x)}{\bar{F}_j(x)} &= \frac{\exp\left(-\frac{1}{2}(\frac{\log(x)-\mu}{\sigma})^2\right)}{\sqrt{2\pi}(\frac{\log(x)-\mu}{\sigma})x^{\frac{-1}{\xi_j}}L(x)} \\ &= \frac{\exp\left(\frac{-1}{2}(\frac{\log(x)-\mu}{\sigma})^2 + (\alpha+\frac{1}{\xi_j})\log(x)\right)}{\sqrt{2\pi}(\frac{\log(x)-\mu}{\sigma})x^{\alpha}L(x)} \to_{x\to\infty} 0\end{aligned}$$

and

$$\frac{\bar{F}(x)}{\bar{F}_j(x)} = \sum_{i\neq j}^{m} p_i \frac{\bar{F}_i(x)}{\bar{F}_j(x)} + p_j,$$

then we can conclude that $0 < \lim_{x\to\infty} \frac{\bar{F}(x)}{\bar{F}_j(x)} = p_j < \infty$*, then* $r(F) = r(F_j)$ *and both distributions have the same MDA.*

Theorem 4 *Let* F *be a cdf that is expressed as* $F(x) = \sum_{i=1}^{m} p_i F_i(x)$*, with* $\sum_{i=1}^{m} p_i = 1,\ \forall p_i > 0$*, if* $j \in \{1, \ldots, m\}$ *is such that* $F_j \in MDA(G_{\xi_j})$*, with* $\xi_j > 0$ *(Fréchet), and* $F_i\ \forall i \neq j$ *have* $MDA(G_{\xi_i})$*, with* $\xi_i = 0$ *(Gumbel), then* $F \in MDA(G_{\xi_j})$*.*

Proof 4 **Case 1:** *If* $r(F_i) = \infty,\ \forall i \neq j$*, and we can find* $\alpha > 0$ *such that* $\frac{1}{\xi_j} + \alpha > 0$*, we obtain:*

$$\frac{\bar{F}_i(x)}{\bar{F}_j(x)} = \frac{\bar{F}_i(x)}{x^{\frac{-1}{\xi_j}}L(x)} = \frac{\bar{F}_i(x)}{x^{-(\alpha+\frac{1}{\xi_j})}}\frac{1}{x^{\alpha}L(x)}.$$

from the properties of the von Mises functions , $\bar{F}_i$ *decreases to zero much faster than* $x^{-\alpha}$*, then when* $r(F_i) = \infty$ *we have (see, Embrechts et al., 1997, page 139):*

$$\lim_{x\to\infty} \frac{\bar{F}_i(x)}{x^{-(\alpha+\frac{1}{\xi_j})}} = 0 \quad and \quad \lim_{x\to\infty} \frac{1}{x^{\alpha}L(x)} = 0,$$

and we conclude that F *and* F_j *are tail equivalents.*

Case 2: *If* $l \neq i \neq j$ *is such that* $r(F_l) < \infty$

$$\frac{\bar{F}(x)}{\bar{F}_j(x)} = p_l \frac{\bar{F}_l(x)}{\bar{F}_j(x)} + \sum_{i\neq l\neq j}^{m} \frac{\bar{F}_i(x)}{\bar{F}_j(x)} + p_j.$$

Let X_l be a random variable with probability distribution function (pdf) $f_l(\cdot)$ with $E(X_l^k) < \infty$ for every $k > 0$,

$$\frac{\bar{F}_l(x)}{\bar{F}_j(x)} = \frac{\bar{F}_l(x)}{(x - r(F_l))f_l(x)} \frac{(x - r(F_l))f_l(x)}{x^{\frac{-1}{\xi_j}} L(x)},$$

the limits of the first term are (see, Embrechts et al., 1997; McNeil et al., 2005):

$$\lim_{x\to\infty} \frac{\bar{F}_l(x)}{(x - r(F_l))f_l(x)} = \lim_{r(F)\to\infty} \lim_{x\to r(F_l)} \frac{\bar{F}_l(x)}{(x - r(F_l))f_l(x)} = 0.$$

We obtain:

$$\frac{(x - r(F_l))f_l(x)}{x^{\frac{-1}{\xi_j}} L(x)} = \frac{x^{\frac{1}{\xi_j}+\alpha}(x - r(F_l))f_l(x)}{x^{\alpha} L(x)} \sim \frac{x^a f_l(x)}{x^{\alpha} L(x)} \to 0, \text{ with } a > 1 \text{ and } \alpha > 0$$

and we achieve the same results as in Case 1.

3 Classical kernel estimator with bias reducing technique

The BCCKE proposed by Kim et al. (2006) can be expressed as a linear combination of the CKE of the pdf, f_X, and the CKE of the cdf, F_X. Let us assume that X_i, $i = 1, ..., n$, denotes data observations from the random variable X; the usual expression for the classical kernel estimator of the pdf is (see, Silverman, 1986):

$$\hat{f}_X(x) = \frac{1}{nb} \sum_{i=1}^{n} k\left(\frac{x - X_i}{b}\right)$$

and for the cdf is (see, Azzalini, 1981):

$$\widehat{F}_X(x) = \int_{-\infty}^{x} \widehat{f}_X(u)du = \tfrac{1}{n} \sum_{i=1}^{n} K\left(\tfrac{x - X_i}{b}\right),$$

where $K(\cdot)$ is the cdf associated with $k(\cdot)$ which is known as the kernel function (usually a bounded and symmetric pdf). Some examples of very common kernel functions are the Epanechnikov and the Gaussian kernel. The parameter b is the bandwidth or the smoothing parameter and it controls the smoothness of the resulting estimation. The larger the value of b, the smoother the resulting estimated function. In practice, the value of b depends on the sample size and satisfies the condition if $n \to \infty$, $b \to 0$ and $nb \to \infty$.

The BCCKE is:

$$\widetilde{F}_X(x) = \frac{\lambda_1 \widehat{F}_1(x) + \widehat{F}_X(x) + \lambda_2 \widehat{F}_2(x)}{\lambda_1 + 1 + \lambda_2}, \tag{6}$$

where $\lambda_1, \lambda_2 > 0$ are weights and

$$\widehat{F}_j(x) = \widehat{F}_X(x + l_j b) - l_j b \widehat{f}_X(x + l_j b), \quad j = 1, 2.$$

Kim et al. (2006) proved that if $\lambda_1 = \lambda_2 = \lambda$ then $-l_1 = l_2 = l(\lambda)$, being:

$$l(\lambda) = \left(\frac{(1 + 2\lambda)\mu_2}{2\lambda} \right)^{1/2},$$

where $\mu_p = \int t^p k(t) dt$. Kim et al. (2006) also proved that the bias of $\widetilde{F}_X(x)$ is $O(b^4)$, while the bias of $\widehat{F}_X(x)$ is $O(b^2)$.

The mean integrated squared error (MISE) can be expressed as the sum of the integrated variance and the integrated square bias:

$$MISE\left(\widetilde{F}_X\right) = E\left(\int (\widetilde{F}_X(x) - F_X(x))^2 dx\right) = \int Var\left(\widetilde{F}_X(x)\right) dx + \int \left[Bias\left(\widetilde{F}_X(x)\right)\right]^2 dx.$$

Based on the asymptotic expression for bias and variance of BCCKE deduced by Kim et al. (2006) we obtain that the asymptotic mean integrated squared error (A-MISE) is:

$$\begin{aligned} A - MISE\left(\widetilde{F}_X(x)\right) &= \frac{1}{n} \frac{2\lambda^2 + 1}{(2\lambda + 1)^2} \int F_X(x)(1 - F_X(x)) dx + \frac{b}{n} V(\lambda) \\ &+ \frac{b^8}{576} \left(\mu_4 - \frac{3(1 + 6\lambda)\mu_2^2}{2\lambda} \right)^2 \int \left(f_X'''(x) \right)^2 dx, \end{aligned} \tag{7}$$

where, if the kernel k has a compact support $[-1, 1]$, it is obtained that:

$$\begin{aligned} V(\lambda) &= \frac{1}{(2\lambda + 1)^2} \left[(2\lambda^2 + 1) \left(\int_{-1}^{1} k^2(t) dt + l \int_{-1}^{1} K^2(t) dt - 1 \right) \right. \\ &+ 2\lambda \left(\int_{-1}^{1-l} k(t - l) k(t) dt + \int_{-1+l}^{1} k(t) k(t + l) dt + \int_{1-l}^{1} (k(t) + \lambda k(t + l)) dt \right. \\ &- \left. \left. \int_{-1}^{-1+2l} K(t) dt + \lambda \int_{-1+l}^{1-l} (k(t - l) k(t + l) - l^2 K(t - l) K(t + l)) dt \right) \right]. \end{aligned} \tag{8}$$

There exists a value of λ that minimises $V(\lambda)$, and this depends on the selected kernel, if the Epanchnikov kernel is used $\lambda = 0.0799$ and $V(0.0799) = -0.1472244$.

Remark 1 *Let F_X be a cdf with four bounded and continuous derivatives, the optimal bandwidth that minimises A-MISE is:*

$$b^{MISE} = n^{-1/7} \left(\frac{-V(\lambda)}{\frac{\int (f_X'''(t))^2 dt}{72} \left(\mu_4 - \frac{3(1+6\lambda)\mu_2^2}{2\lambda} \right)^2} \right)^{1/7} . \qquad (9)$$

Kim et al. (2006) did not analyse a method to estimate the optimal bandwidth. Similarly to the CKE, we can use iterative methods such as the plug-in methods or the methods based on cross-validation (see, Jones et al., 1996, for a review). Alternatively, the rule-of-thumb bandwidth is a direct way to estimate the smoothing parameter. Following Silverman (1986), for the BCCKE the rule-of-thumb bandwidth is obtained by replacing in expression (9) the functional $\int (f_X'''(x))^2 dx$ with its value assuming that f_X is the density of a normal distribution with scale parameter σ, then:

$$b^{*} = n^{-1/7} \sigma \left(\frac{-V(\lambda)}{\frac{0.5289277}{72} \left(\mu_4 - \frac{3(1+6\lambda)\mu_2^2}{2\lambda} \right)^2} \right)^{1/7} . \qquad (10)$$

The smoothing parameter in (10) can be estimated by replacing σ with a consistent estimation, such as the sample standard deviation s. However, Silverman (1986) noted that for long-tailed and right-skewed distributions it is better to use a robust estimation of σ based on the interquartile range R, that is $\frac{R}{1.34}$. In general, we can use the better estimator of σ for each case: $\hat{\sigma} = Min\left(s, \frac{R}{1.34}\right)$.

4 Transformed kernel estimator with bias reducing technique

In this section we propose a new kernel estimator that combines the greater efficiency of the transformed kernel estimator of the cdf with the bias reduction technique. In general, the transformed kernel estimator involves selecting a transformation function so that the cdf or the pdf associated with the transformed variable can be estimated optimally with the classical kernel estimator or a bias-corrected version. We denote $T(\cdot)$ the transformation function, then the transformed random variable is $Y = T(X)$, and we know that $f_X(x) = f_Y(y)T'(x)$ and $F_X(x) = F_Y(y)$.

Let $T(\cdot)$ be a concave transformation function with at least four continuous derivatives. Assuming equal weights in (6), i.e. $\lambda_1 = \lambda_2 = \lambda > 0$, the bias corrected transformed kernel estimator (BCTKE) is:

$$\widetilde{F}_{T(X)}(T(x)) = \frac{\lambda\left[\widehat{F}_1(T(x)) + \widehat{F}_2(T(x))\right] + \widehat{F}_{T(X)}(T(x))}{2\lambda+1} = \widetilde{\widetilde{F}}_X(x). \tag{11}$$

We denote $y = T(x)$ and the transformed data are $Y_i = T(X_i)$, $i = 1, ..., n$, then:

$$\widehat{F}_{T(x)}(T(x)) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{T(x) - T(X_i)}{b}\right) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{y - Y_i}{b}\right) = \widehat{F}_Y(y) = \widehat{\widehat{F}}_X(x) \tag{12}$$

and

$$\begin{aligned} \widehat{F}_1(T(x)) &= \widehat{F}_{T(X)}(T(x) - lb) + lb\hat{\hat{f}}_X(x - lb) = \widehat{\widehat{F}}_1(x), \\ \widehat{F}_2(T(x)) &= \widehat{F}_{T(X)}(T(x) + lb) - lb\hat{\hat{f}}_X(x + lb) = \widehat{\widehat{F}}_1(x), \end{aligned} \tag{13}$$

where $\widehat{\widehat{f}}_X$ is the transformed kernel density estimation (see, for example, Wand et al., 1991; Buch-Larsen et al., 2005; Bolancé et al., 2008; Bolancé, 2010).

$$\hat{\hat{f}}_X(x) = \frac{1}{nb}\sum_{i=1}^{n} k\left(\frac{T(x) - T(X_i))}{b}\right) T'(x). \tag{14}$$

Theorem 5 *Let F_X be a cdf with four bounded and continuous derivatives. Let $T(\cdot)$ be a concave transformation function with at least four continuous derivatives.. If the kernel k has a compact support $[-1, 1]$, we obtain that the bias and variance of BCTKE are:*

$$E\left(\widetilde{\widetilde{F}}_X(x) - F_X(x)\right) = \frac{b^4}{24}\left(\mu_4 - \frac{3(1+6\lambda)\mu_2^2}{2\lambda}\right)\frac{f_X'''(x)}{T'(x)} D\left(T^{(p)}(x), F_X^{(p)}(x)\right) + o\left(b^4\right), \tag{15}$$

$$Var\left(\widetilde{\widetilde{F}}_X(x)\right) = \frac{1}{n}\frac{2\lambda^2+1}{(2\lambda+1)^2} F_X(x)\left(1 - F_X(x)\right) + \frac{f_X(x)}{T'(x)}\frac{b}{n} V(\lambda) + o\left(\frac{b^2}{n}\right). \tag{16}$$

The function $D\left(T^{(p)}(x), F_X^{(p)}(x)\right)$ with $p = 0, ..., 4$, where the super-index between parentheses refers to the derivative, depends on the transformation T, the cdf F_X and the first four derivatives of these functions, is such that:

$$D\left(T^{(p)}(x), F_X^{(p)}(x)\right) = 0 \;\; \text{if} \;\; T(x) = F(x)$$

and

$$D\left(T^{(p)}(x), F_X^{(p)}(x)\right) \to 0 \;\; \text{if} \;\; T^{(p)}(x) \to F_X^{(p)}(x), \;\; \forall p = 0, ..., 4.$$

Proof 5 *The bias and the variance of the BCTKE are obtained from the bias and variance of the BCCKE of* $\widetilde{F}_Y(y)$*, knowing that* $F_Y(y) = F_X(x)$ *and* $f_Y(y) = \frac{f_X(x)}{T'(x)}$ *and analysing the derivative* $\left(\frac{f_X(x)}{T'(x)}\right)'''$.

$$
\begin{aligned}
\left(\frac{f_X(x)}{T'(x)}\right)''' &= \frac{f_X'''(x)}{T'(x)} - \frac{3f_X''(x)T''(x)}{T'(x)^2} - \frac{3T'''(x)f_X'(x)}{T'(x)^2} + \frac{6f_X'(x)T''(x)^2}{T'(x)^3} - \frac{f_X(x)T^{(4)}(x)}{T'(x)^2} \\
&- \frac{6f_X(x)T''(x)^3}{T'(x)^4} + \frac{6f_X(x)T'''(x)T''(x)}{T'(x)^3},
\end{aligned}
$$

if $T^{(p)}(x) \to F_X^{(p)}(x),\ \forall p = 0, ..., 4$ *we obtain that* $D\left(T^{(p)}(x), F_X^{(p)}(x)\right) \to 0$.

From the results of Theorem 5 we prove that if a suitable transformation is found, we can reduce the bias and the variance of the BCCKE.

4.1 Double transformed kernel estimator with bias correction

The BCDTKE estimator is obtained in a similar manner to that used to obtain the DTKE estimator (see, Alemany et al., 2013).

Let F be a continuous cdf with four bounded and continuous derivatives in a neighbourhood of x, we assume that k is the kernel that is a symmetric pdf and with a compact support $[-1, 1]$ and b is the bandwidth. The smoothing parameter b holds that when $n \to \infty$, $b \to 0$ and $nb \to \infty$, then the A-MISE associated with the BDCKE of the transformed random variable Y is:

$$
\begin{aligned}
&\frac{1}{n}\frac{2\lambda^2+1}{(2\lambda+1)^2}\int F_Y(y)(1-F_Y(y))dx + \frac{b}{n}V(\lambda) \\
&\quad + \frac{b^8}{576}\left(\mu_4 - \frac{3(1+6\lambda)\mu_2^2}{2\lambda}\right)^2 \int (f_Y'''(y))^2 dx
\end{aligned}
$$

where $V(\lambda) < 0$ is the function defined in (8).

Given b and the kernel k, the A-MISE is minimum when functional $\int [f_Y'''(y)]^2\, dy$ is minimum. The proposed method is based on the transformation of the variable in order to achieve a distribution that minimises the A-MISE, i.e. that minimises $\int [f_Y'''(y)]^2\, dy$.

Terrell (1990) showed that the density of a $Beta\,(5, 5)$ distribution defined on the domain $[-1, 1]$ minimises $\int [f_Y'''(y)]^2\, dy$, in the set of all densities with known variance. The pdf h and cdf H of the $Beta\,(5,5)$ are:

$$\begin{aligned} h(x) &= \frac{315}{256}(1-x^2)^4 \quad -1 \leq x \leq 1, \\ H(x) &= \frac{1}{256}(35x^4 - 175x^3 + 345x^2 - 325x + 128)(x+1)^5. \end{aligned}$$

Then the BCDTKE is:

$$\widetilde{F}_{H^{-1}(T(X))}(H^{-1}\left(T\left(x\right)\right)) =$$

$$= \frac{\lambda\left[\widehat{F}_{\{H^{-1}(T(X)),1\}}(H^{-1}\left(T\left(x\right)\right)) + \widehat{F}_{\{H^{-1}(T(x)),2\}}(H^{-1}\left(T\left(x\right)\right))\right] + \widehat{F}_{H^{-1}(T(x))}(H^{-1}\left(T\left(x\right)\right))}{2\lambda+1} = \widetilde{\widetilde{F}}_X(x)$$

where $T(.)$ is a first transformation that matches a cdf, so that the transformed sample $T(X_i)$, $i = 1, ..., n$, takes values from a $Uniform(0,1)$ distribution and, therefore, the double transformed sample $H^{-1}(T(X_i))$, $i = 1, ..., n$, takes values from a $Beta\,(5,5)$ distribution. Similarly to (13), we obtain that

$$\begin{aligned} \widehat{F}_{\{H^{-1}(T(x)),1\}}(x) &= \widehat{F}_{H^{-1}(T(x))}H^{-1}\left(T\left(x-lb\right)\right) + lb\widehat{f}_{H^{-1}(T(x))}H^{-1}\left(T\left(x-lb\right)\right), \\ \widehat{F}_{\{H^{-1}(T(x)),2\}}(x) &= \widehat{F}_{H^{-1}(T(x))}H^{-1}\left(T\left(x+lb\right)\right) - lb\widehat{f}_{H^{-1}(T(x))}H^{-1}\left(T\left(x+lb\right)\right), \end{aligned}$$

where $\widehat{f}_{H^{-1}(T(x))}$ is the double transformed kernel density estimation (see, Bolancé et al., 2008; Bolancé, 2010):

$$\begin{aligned} &\widehat{f}_{H^{-1}(T(X))}(H^{-1}\left(T\left(x\right)\right)) = \\ &\frac{1}{nb}\sum_{i=1}^{n} k\left(\frac{H^{-1}\left(T\left(x\right)\right) - H^{-1}\left(T\left(X_i\right)\right)}{b}\right) H^{-1'}\left(T\left(x\right)\right)T'(x). \end{aligned}$$

The smoothing parameter b in BCDTKE can be calculated from expressions (9) replacing f''' by $Beta(5,5)$ pdf:

$$b^* = n^{-1/7}\left(\frac{-V(\lambda)}{\frac{1288.6}{72}\left(\mu_4 - \frac{3(1+6\lambda)\mu_2^2}{2\lambda}\right)^2}\right)^{1/7}. \qquad (17)$$

5 Simulation study

We compare four kernel estimation methods: CKE, BCCKE, DTKE and BCDTKE. The first transformation $T(\cdot)$ that we use for obtaining DTKE and BCDTKE is the cdf of the modified Champernowne distribution[2] analysed by Buch-Larsen et al. (2005). These authors also proposed a method based on maximising a pseudo-likelihood function to estimate the parameters. We use the rule-of-thumb bandwidth based on minimising A-MISE.

To compare estimated cdfs with theoretical cdfs we use two distances:

$$\begin{aligned} L_1(\check{F}) &= \int \left|\check{F}(t) - F(t)\right| dt \\ L_2(\check{F}) &= \int \left(\check{F}(t) - F(t)\right)^2 dt, \end{aligned} \tag{18}$$

where $\check{F}$ represents the different estimators. Distances L_1 and L_2 evaluate the fit of the cdf differently. Distance L_2 attaches greater importance to the major differences between the theoretical cdf and the fitted cdf than is attached by distance L_1. When the aim is to fit an extreme value distribution, estimation errors tend to increase as the cdf approaches 1, due to a lack of sample information on the extreme values of the variable. Therefore, distance L_2 will be more strongly influenced by the estimation errors at the extreme values of the variable.

We generated 2000 samples for each sample size analysed: $n = 100$, $n = 500$, $n = 1000$ and $n = 5000$ and for each distribution in Table 1. We selected four distributions[3] that are positively skewed and which present different tail shapes: Lognormal, Weibull (both Gumbel types) and two mixtures of Lognormal-Pareto (both Fréchet types). The Lognormal and Weibull distributions both have an exponential tail. Specifically, we define the Weibull distribution with a scale parameter equal to 1 and shape parameter γ, so that the smaller the value of γ the slower is the exponential decay in the tail, i.e. the lower the value of γ, the lighter the tail. For the Lognormal distribution, the shape parameter is σ. In this case, the higher the value of σ, the lighter the tail. Furthermore, we analyse two mixtures of Lognormal-Pareto, that is, distributions with "fat" tails or heavy-tailed distributions. As we proved in section 2, these mixtures are Fréchet type and have a Pareto tail; thus, in this case the smaller the value of shape parameter ρ, the heavier is the tail.

[2]The cdf of the modified Champernowne distribution is:

$$T(x) = \frac{(x+c)^{\alpha} - c^{\alpha}}{(x+c)^{\alpha} + (M+c)^{\alpha} - 2c^{\alpha}}, \; for\, x \geq 0, \; \alpha, M, c > 0.$$

[3]We used the same parameters as in Alemany et al. (2013, 2012).

For each simulated sample, we estimated the cdf using the four methods: CKE, BCCKE, DTKE and BCDTKE and we calculated the distances defined in (18). Finally, for each sample size, we calculated the mean of the 2000 replicates. To calculate distances L_1 and L_2 with each simulated sample we used the grid proposed by Buch-Larsen et al. (2005) based on the change of variable defined by Clements et al. (2003), $y = \frac{x-M}{x+M}$, where M is the sample median.

To obtain CKE and BCCKE we used two smoothing parameters: the rule-of-thumb, estimating σ from the sample standard deviation s and from $Min\left(s, \frac{R}{1.34}\right)$, where R is the sample interquartile range. The results obtained with s are shown in Tables 4 and 5 in the Appendix. Specifically, from the results in Table 5, we can conclude that both estimators $-$CKE and BCCKE using rule-of-thumb, estimating σ from the sample standard deviation $s-$ are not consistent when the distribution is heavy tailed.

Table 1: Distributions in the simulation study

Distribution	$F_X(x)$	Parameters
Weibull	$1 - e^{-x^\gamma}$	$\gamma = 0.75$
		$\gamma = 1.5$
		$\gamma = 3$
Lognormal	$\int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$	$(\mu, \sigma = (0, 0.25)$
		$(\mu, \sigma) = (0, 0.5)$
		$(\mu, \sigma) = (0, 1.0)$
Mixture of Lognormal -Pareto	$p \int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt + (1-p)\left(1 - \left(\frac{x-c}{\lambda}\right)^{-\rho}\right)$	$(p, \mu, \sigma, \lambda, \rho, c) = (0.7, 0, 1, 1, 0.9, -1)$
		$(p, \mu, \sigma, \lambda, \rho, c) = (0.3, 0, 1, 1, 0.9, -1)$
		$(p, \mu, \sigma, \lambda, \rho, c) = (0.7, 0, 1, 1, 1.0, -1)$
		$(p, \mu, \sigma, \lambda, \rho, c) = (0.3, 0, 1, 1, 1.0, -1)$
		$(p, \mu, \sigma, \lambda, \rho, c) = (0.7, 0, 1, 1, 1.1, -1)$
		$(p, \mu, \sigma, \lambda, \rho, c) = (0.3, 0, 1, 1, 1.1, -1)$

In Tables 2 and 3 we compare the BCCKE, the DTKE and the BCDTKE with the CKE, i.e., we obtain the ratio between distances L_1 and L_2 that were obtained with the BCCKE, the DTKE and the BCDTKE and those that were obtained with the CKE. If the ratio is greater than 1, then the CKE is better; if it is lower, then the corrected estimator improves the CKE. The absolute distances are shown in Tables 6 and 7 in the Appendix.

The results presented in Tables 2 and 3 point to differences between distances L_1 and L_2 and, furthermore, there exist important differences between the results obtained for Gumbel-type and Fréchet-type distributions.

Focusing first on the DTKE, for distance L_1 this estimator does not improve the CKE in any case. Furthermore, when the sample is small the L_1 obtained for the DTKE is considerably worse than that obtained for the CKE. For distance L_2 the DTKE improves the CKE in small and large sample sizes. Focusing on L_2, we observe that the largest improvements of the DTKE occur when the distributions are Fréchet-type, although these improvements are not as great as those obtained when bias correction is used.

Focusing now on Gumbel-type distributions, the results in Table 2 show that, in general, both boundary correction approaches, the BCCKE and the BCDTKE, make similar improvements to the CKE in distance L_2 for all sample sizes. Furthermore, the improvement is greater as the sample size increases. For distance L_1 the BCCKE and the BCDTKE do not improve the CKE when the distribution has a lighter tail, i.e., the Weibull distributions with larger shape parameter and the Lognormal distributions with smaller shape parameter.

In Table 3 we show the results for the Fréchet-type distributions. We observe that, when the distribution has a heavier tail, the improvement of the BCDTKE with respect to the CKE is more marked than that obtained with BCCKE, for all sample sizes and both distances, except for distance L_1 in the case of 70Lognormal-30Pareto ($\rho = 1.1$) and sample size 100. In general, for distance L_2 the improvement of the BCDTKE with respect to the BCCKE is around 5%. For distance L_1 this improvement becomes greater as the sample size increases, exceeding 10% in the case of 70Lognormal-30Pareto ($\rho = 1$).

6 Conclusions

In many analyses −be it in economics, finance, insurance, demography, etc.− the fit of the cdf is of great interest for evaluating the probability of extreme situations. In such cases, the data are usually generated by a continuous random variable X whose distribution may result from the mixture of different EVDs; however, in such instances both classical parametric models and classical nonparametric estimates cannot be used to estimate the cdf. We have presented a method for estimating the cdf that is suitable when the loss (or whatever the analysed variable may be) is a heavy-tailed random variable. The double transformation kernel using bias-corrected technique proposed here provides, in general, a good fit for Gumbel and Fréchet extreme value distribution types, especially when the sample size is small.

We show, for a small sample size, that the bias-corrected double transformed kernel estimator proposed here improves the classical kernel estimator and bias-corrected classical kernel estimator of cumulative distribution function when the distribution is an extreme value distribution and the maximum domain of attraction is the associated with a type Fréchet-type distribution.

Table 2: Comparative ratios obtained with the simulation results for Weibull and Lognormal (Gumbel-type distributions) using rule-of-thumb with scale parameter $Min\left(s, \frac{R}{1.34}\right)$.

n	100		500		1000		5000	
				Lognormal ($\sigma = 0.25$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	1.0312	0.2002	1.0275	0.1334	1.0238	0.1112	1.0136	0.0742
DTKE	297.5476	0.6184	37.9005	0.1506	13.2495	0.1127	1.7839	0.0734
BCDTKE	1.0361	0.2030	1.0307	0.1350	1.0264	0.1124	1.0153	0.0748
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9777	0.1882	0.9789	0.1235	0.9830	0.1051	0.9892	0.0703
DTKE	115.4618	0.4155	16.0135	0.1331	6.1701	0.1062	1.4138	0.0701
BCDTKE	0.9680	0.1885	0.9693	0.1236	0.9738	0.1052	0.9811	0.0704
				Lognormal ($\sigma = 1$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9486	0.1625	0.9433	0.1069	0.9416	0.0883	0.9505	0.0588
DTKE	43.4451	0.3195	7.4330	0.1213	3.8088	0.0944	1.4983	0.0604
BCDTKE	0.9194	0.1598	0.9137	0.1054	0.9112	0.0871	0.9230	0.0582
				Weibull ($\gamma = 0.75$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9626	0.1768	0.9444	0.1169	0.9454	0.0988	0.9383	0.0661
DTKE	15.5507	0.2372	2.1968	0.1254	1.4867	0.1024	1.1714	0.0657
BCDTKE	0.9338	0.1740	0.9139	0.1147	0.9140	0.0965	0.9015	0.0630
				Weibull ($\gamma = 1.5$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	1.0148	0.1996	0.9874	0.1321	0.9828	0.1114	0.9762	0.0731
DTKE	55.5021	0.2919	5.9624	0.1322	2.1170	0.1094	1.0632	0.0725
BCDTKE	1.0121	0.2006	0.9849	0.1327	0.9801	0.1119	0.9733	0.0733
				Weibull ($\gamma = 3$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	1.0644	0.2103	1.0396	0.1384	1.0328	0.1155	1.0168	0.0761
DTKE	53.9094	0.2656	2.4343	0.1357	1.3076	0.1134	1.0787	0.0755
BCDTKE	1.0699	0.2126	1.0440	0.1397	1.0365	0.1165	1.0200	0.0766

Table 3: Comparative ratios obtained with the simulation results for Mixtures of Lognormal-Pareto (Fréchet-type distributions) using rule-of-thumb with scale parameter $Min\left(s, \frac{R}{1.34}\right)$.

n	100		500		1000		5000	
	70Lognormal-30Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9972	0.0767	0.9981	0.0434	0.9983	0.0342	0.9986	0.0260
DTKE	7.1804	0.2000	3.6744	0.0878	2.9466	0.0637	2.4752	0.0441
BCDTKE	0.9656	0.0724	0.9377	0.0411	0.9121	0.0322	0.9283	0.0247
	70Lognormal-30Pareto ($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9950	0.0851	0.9970	0.0463	0.9975	0.0360	0.9982	0.0209
DTKE	10.3436	0.2193	4.6247	0.0895	3.3365	0.0614	2.3266	0.0318
BCDTKE	0.9948	0.0814	0.9490	0.0441	0.9259	0.0342	0.8928	0.0199
	70Lognormal-30Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9930	0.0943	0.9953	0.0519	0.9960	0.0401	0.9975	0.0222
DTKE	14.2912	0.2441	6.0630	0.0954	4.4212	0.0655	2.3962	0.0301
BCDTKE	1.0007	0.0908	0.9650	0.0499	0.9512	0.0386	0.9164	0.0214
	30Lognormal-70Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9975	0.0804	0.9982	0.0464	0.9984	0.0373	0.9986	0.0264
DTKE	4.6250	0.1790	2.3205	0.0757	2.0276	0.0571	1.7288	0.0347
BCDTKE	0.9698	0.0759	0.9123	0.0435	0.9084	0.0351	0.9479	0.0252
	30Lognormal-70Pareto($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9972	0.0842	0.9976	0.0476	0.9980	0.0373	0.9983	0.0249
DTKE	6.2399	0.1963	2.9780	0.0794	2.3838	0.0570	1.7963	0.0319
BCDTKE	0.9619	0.0794	0.9227	0.0448	0.9182	0.0353	0.9360	0.0237
	30Lognormal-70Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BCCKE	0.9958	0.0911	0.9967	0.0508	0.9973	0.0391	0.9982	0.0233
DTKE	8.6851	0.2189	3.8710	0.0860	2.9610	0.0604	1.7962	0.0296
BCDTKE	0.9716	0.0867	0.9472	0.0483	0.9288	0.0372	0.9011	0.0223

Appendix

Table 4: Simulation results for Weibull and Lognormal using rule-of-thumb with scale parameter s.

n	100		500		1000		5000	
				Lognormal ($\sigma = 0.25$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0161	0.1239	0.0073	0.0838	0.0051	0.0704	0.0023	0.0476
BCCKE	0.0163	0.0246	0.0075	0.0111	0.0052	0.0078	0.0024	0.0035
				Lognormal ($\sigma = 0.5$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0410	0.1983	0.0180	0.1321	0.0131	0.1126	0.0058	0.0747
BCCKE	0.0378	0.0361	0.0169	0.0159	0.0124	0.0116	0.0055	0.0052
				Lognormal ($\sigma = 1$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.1735	0.4069	0.0848	0.2872	0.0611	0.2442	0.0273	0.1635
BCCKE	0.1338	0.0592	0.0644	0.0273	0.0462	0.0192	0.0214	0.0087
				Weibull ($\gamma = 0.75$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.1116	0.3265	0.0534	0.2276	0.0387	0.1940	0.0179	0.1321
BCCKE	0.0966	0.0545	0.0452	0.0253	0.0328	0.0185	0.0151	0.0088
				Weibull ($\gamma = 1.5$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0360	0.1854	0.0170	0.1278	0.0122	0.1085	0.0054	0.0724
BCCKE	0.0362	0.0368	0.0167	0.0168	0.0120	0.0121	0.0053	0.0053
				Weibull ($\gamma = 3$)				
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0187	0.1330	0.0085	0.0902	0.0061	0.0766	0.0027	0.0515
BCCKE	0.0198	0.0280	0.0088	0.0125	0.0063	0.0088	0.0028	0.0039

Table 5: Simulation results for Mixtures of Lognormal-Pareto using rule-of-thumb with scale parameter s.

n	100		500		1000		5000	
	70Lognormal-30Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	8.0598	1.9372	9.4548	2.0452	12.6591	2.1364	13.2365	2.3089
CKEbrt	5.0740	0.2238	5.0508	0.2990	6.8743	0.3847	7.4525	0.6335
	70Lognormal-30Pareto ($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	2.9700	1.4059	3.4953	1.4343	5.0217	1.4872	5.6733	1.4808
CKEbrt	1.9101	0.1472	1.9262	0.1721	2.4518	0.2140	3.4110	0.3214
	70Lognormal-30Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	2.5304	1.0949	1.7559	1.0465	2.8612	1.0500	2.3528	1.0210
CKEbrt	1.2847	0.1156	1.0331	0.1066	1.9549	0.1304	1.3763	0.1758
	30Lognormal-70Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	18.5349	2.9764	23.3374	3.2261	25.2348	3.2729	30.4567	3.4660
CKEbrt	11.2088	0.3977	12.3226	0.5722	12.5870	0.6789	19.1204	1.0613
	30Lognormal-70Pareto($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	10.6396	2.1808	6.2217	2.0786	9.6786	2.1426	9.5012	2.0563
CKEbrt	5.6933	0.2687	3.5953	0.2862	5.0031	0.3613	5.6010	0.5123
	30Lognormal-70Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	4.9616	1.5990	4.8501	1.5662	4.5138	1.5305	4.0134	1.4511
CKEbrt	2.5866	0.1808	3.1662	0.2080	2.8573	0.2260	2.1638	0.2942

Table 6: Simulation results for Weibull and Lognormal using rule-of-thumb with scale parameter $Min\left(s, \frac{R}{1.34}\right)$.

n	100		500		1000		5000	
	Lognormal ($\sigma = 0.25$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0159	0.1232	0.0073	0.0837	0.0051	0.0703	0.0023	0.0475
BCCKE	0.0164	0.0247	0.0075	0.0112	0.0052	0.0078	0.0024	0.0035
DTKE	4.7257	0.0762	0.2756	0.0126	0.0678	0.0079	0.0042	0.0035
BCDTKE	0.0165	0.0250	0.0075	0.0113	0.0053	0.0079	0.0024	0.0036
	Lognormal ($\sigma = 0.5$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0389	0.1931	0.0173	0.1294	0.0126	0.1107	0.0056	0.0739
BCCKE	0.0380	0.0363	0.0170	0.0160	0.0124	0.0116	0.0056	0.0052
DTKE	4.4878	0.0802	0.2776	0.0172	0.0780	0.0118	0.0079	0.0052
BCDTKE	0.0376	0.0364	0.0168	0.0160	0.0123	0.0116	0.0055	0.0052
	Lognormal ($\sigma = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.1415	0.3695	0.0680	0.2575	0.0488	0.2182	0.0222	0.1472
BCCKE	0.1342	0.0600	0.0642	0.0275	0.0460	0.0193	0.0211	0.0087
DTKE	6.1486	0.1180	0.5058	0.0312	0.1859	0.0206	0.0333	0.0089
BCDTKE	0.1301	0.0591	0.0622	0.0272	0.0445	0.0190	0.0205	0.0086
	Weibull ($\gamma = 0.75$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.1008	0.3107	0.0476	0.2147	0.0344	0.1829	0.0157	0.1236
BCCKE	0.0970	0.0549	0.0450	0.0251	0.0326	0.0181	0.0147	0.0082
DTKE	1.5673	0.0737	0.1046	0.0269	0.0512	0.0187	0.0184	0.0081
BCDTKE	0.0941	0.0541	0.0435	0.0246	0.0315	0.0176	0.0142	0.0078
	Weibull ($\gamma = 1.5$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0357	0.1848	0.0169	0.1276	0.0122	0.1083	0.0054	0.0723
BCCKE	0.0363	0.0369	0.0167	0.0169	0.0120	0.0121	0.0053	0.0053
DTKE	1.9834	0.0540	0.1009	0.0169	0.0258	0.0119	0.0057	0.0052
BCDTKE	0.0362	0.0371	0.0167	0.0169	0.0119	0.0121	0.0053	0.0053
	Weibull ($\gamma = 3$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	0.0187	0.1330	0.0085	0.0902	0.0061	0.0766	0.0027	0.0515
BCCKE	0.0199	0.0280	0.0088	0.0125	0.0063	0.0088	0.0028	0.0039
DTKE	1.0059	0.0353	0.0207	0.0122	0.0080	0.0087	0.0030	0.0039
BCDTKE	0.0200	0.0283	0.0089	0.0126	0.0063	0.0089	0.0028	0.0039

Table 7: Simulation results for Mixtures of Lognormal-Pareto using rule-of-thumb with scale parameter $Min\left(s, \frac{R}{1.34}\right)$.

n	100		500		1000		5000	
	70Lognormal-30Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	3.0542	1.5946	2.3282	1.4453	1.9522	1.3488	0.9512	0.9403
BCCKE	3.0457	0.1223	2.3239	0.0628	1.9490	0.0461	0.9498	0.0210
DTKE	21.9300	0.3189	8.5546	0.1269	5.7525	0.0860	2.3543	0.0356
BCDTKE	2.9490	0.1155	2.1832	0.0594	1.7806	0.0434	0.8830	0.0200
	70Lognormal-30Pareto ($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.6856	1.1958	1.3384	1.0788	1.1827	1.0244	0.6770	0.8077
BCCKE	1.6771	0.1018	1.3343	0.0500	1.1797	0.0369	0.6758	0.0169
DTKE	17.4350	0.2622	6.1896	0.0965	3.9460	0.0629	1.5752	0.0257
BCDTKE	1.6768	0.0973	1.2701	0.0476	1.0950	0.0351	0.6045	0.0161
	70Lognormal-30Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	1.0170	0.9261	0.7713	0.8170	0.6550	0.7628	0.4306	0.6341
BCCKE	1.0099	0.0873	0.7677	0.0424	0.6524	0.0306	0.4295	0.0141
DTKE	14.5344	0.2261	4.6766	0.0779	2.8958	0.0500	1.0318	0.0191
BCDTKE	1.0177	0.0841	0.7443	0.0408	0.6230	0.0295	0.3946	0.0136
	30Lognormal-70Pareto ($\rho = 0.9$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	6.3599	2.3361	4.5735	2.0597	3.4532	1.8223	1.5239	1.1857
BCCKE	6.3439	0.1877	4.5654	0.0956	3.4477	0.0679	1.5218	0.0313
DTKE	29.4148	0.4181	10.6127	0.1558	7.0016	0.1041	2.6346	0.0411
BCDTKE	6.1676	0.1774	4.1721	0.0897	3.1368	0.0640	1.4445	0.0298
	30Lognormal-70Pareto($\rho = 1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	3.6494	1.7381	2.5657	1.5209	2.0622	1.3898	0.9779	0.9545
BCCKE	3.6391	0.1464	2.5596	0.0723	2.0582	0.0519	0.9763	0.0238
DTKE	22.7718	0.3411	7.6408	0.1208	4.9160	0.0792	1.7567	0.0305
BCDTKE	3.5104	0.1381	2.3675	0.0682	1.8936	0.0491	0.9153	0.0226
	30Lognormal-70Pareto ($\rho = 1.1$)							
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
CKE	2.0891	1.3123	1.5079	1.1466	1.2428	1.0595	0.7175	0.8321
BCCKE	2.0803	0.1196	1.5029	0.0582	1.2394	0.0415	0.7162	0.0194
DTKE	18.1441	0.2872	5.8372	0.0986	3.6800	0.0640	1.2888	0.0246
BCDTKE	2.0298	0.1138	1.4284	0.0554	1.1543	0.0395	0.6466	0.0185

Acknowledgements

We thank the Spanish Ministry of Science / FEDER support ECO2010-21787-C03-01. We also thank members and affiliates of Riskcenter at the University of Barcelona.

References

Alemany, R., Bolancé, C., Guillén, M., 2012. Non-parametric estimation of value-at-risk. XARXA de Referencia en Economia Aplicada Working paper XREAP2012-19.

Alemany, R., Bolancé, C., Guillén, M., 2013. A nonparametric approach to calculating value-at-risk. Insurance: Mathematics and Economics 52, 255–262.

Azzalini, A., 1981. A note on the estimation of a distribution function and quantiles by a kernel method. Biometrika 68, 326–328.

Bolancé, C., 2010. Optimal inverse beta(3,3) transformation in kernel density estimation. SORT-Statistics and Operations Research Transactions 34, 223–237.

Bolancé, C., Guillén, M., Nielsen, J., 2008. Inverse beta transformation in kernel density estimation. Statistics& Probability Letters 78, 1757–1764.

Buch-Larsen, T., Guillén, M., Nielsen, J., Bolancé, C., 2005. Kernel density estimation for heavy-tailed distributions using the Champernowne transformation. Statistics 39, 503–518.

Choi, E., Hall, P., 1998. On bias reduction in local linear smoothing. Biometrika 85, 333–345.

Clements, A., Hurn, A., Lindsay, K., 2003. Möbius-like mappings and their use in kernel density estimation. Journal of the American Statistical Association 98, 993–1000.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modeling Extremal Events for Insurance and Finance, 4th Edition. Springer-Verlag, Zurich.

Jenkinson, A. F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological events. Quarterly Journal of the Royal Meteorological Society 158–172, 81.

Jones, M., Marron, J., Sheather, S., 1996. A brief survey of bandwidth selection for density estimation. Journal of the American Statistical Association 91, 401–407.

Kim, C., Kim, S., Park, M., Lee, H., 2006. A bias reducing technique in kernel distribution function estimation. Computational Statistics 21, 589–601.

McNeil, A., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press, Princeton.

Reiss, R.-D., Thomas, M., 1997. Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and Others Fields. Birkhäuser Verlag, Berlin.

Silverman, B., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC Finance Series, London.

Terrell, G., 1990. The maximal smoothing principle indensity estimation. Journal of the American Statistical Association 85, 270–277.

Wand, P., Marron, J., Ruppert, D., 1991. Transformations in density estimation. Journal of the American Statistical Association 86, 343–361.

Wang, S., 1995. Insurance pricing and increased limits ratemaking by proportional hazards transforms. Insurance: Mathematics and Economics 17, 43–54.

Wang, S., 1996. Premium calculation by transforming the layer premium density. ASTIN Bulletin 26, 71–92.

2006

CREAP2006-01
Matas, A. (GEAP)**; Raymond, J.Ll.** (GEAP)
"Economic development and changes in car ownership patterns"
(Juny 2006)

CREAP2006-02
Trillas, F. (IEB)**; Montolio, D.** (IEB)**; Duch, N.** (IEB)
"Productive efficiency and regulatory reform: The case of Vehicle Inspection Services"
(Setembre 2006)

CREAP2006-03
Bel, G. (PPRE-IREA)**; Fageda, X.** (PPRE-IREA)
"Factors explaining local privatization: A meta-regression analysis"
(Octubre 2006)

CREAP2006-04
Fernàndez-Villadangos, L. (PPRE-IREA)
"Are two-part tariffs efficient when consumers plan ahead?: An empirical study"
(Octubre 2006)

CREAP2006-05
Artís, M. (AQR-IREA)**; Ramos, R.** (AQR-IREA)**; Suriñach, J.** (AQR-IREA)
"Job losses, outsourcing and relocation: Empirical evidence using microdata"
(Octubre 2006)

CREAP2006-06
Alcañiz, M. (RISC-IREA)**; Costa, A.; Guillén, M.** (RISC-IREA)**; Luna, C.; Rovira, C.**
"Calculation of the variance in surveys of the economic climate"
(Novembre 2006)

CREAP2006-07
Albalate, D. (PPRE-IREA)
"Lowering blood alcohol content levels to save lives: The European Experience"
(Desembre 2006)

CREAP2006-08
Garrido, A. (IEB)**; Arqué, P.** (IEB)
"The choice of banking firm: Are the interest rate a significant criteria?"
(Desembre 2006)

CREAP2006-09
Segarra, A. (GRIT)**; Teruel-Carrizosa, M.** (GRIT)
"Productivity growth and competition in spanish manufacturing firms:
What has happened in recent years?"
(Desembre 2006)

CREAP2006-10
Andonova, V.; Díaz-Serrano, Luis. (CREB)
"Political institutions and the development of telecommunications"
(Desembre 2006)

CREAP2006-11
Raymond, J.L.(GEAP)**; Roig, J.L..** (GEAP)
"Capital humano: un análisis comparativo Catalunya-España"
(Desembre 2006)

CREAP2006-12
Rodríguez, M.(CREB)**; Stoyanova, A.** (CREB)
"Changes in the demand for private medical insurance following a shift in tax incentives"
(Desembre 2006)

CREAP2006-13
Royuela, V. (AQR-IREA)**; Lambiri, D.**; **Biagi**, **B.**
"Economía urbana y calidad de vida. Una revisión del estado del conocimiento en España"
(Desembre 2006)

CREAP2006-14
Camarero, M.; Carrion-i-Silvestre, J.LL. (AQR-IREA)**.**;**Tamarit**, **C.**
"New evidence of the real interest rate parity for OECD countries using panel unit root tests with breaks"
(Desembre 2006)

CREAP2006-15
Karanassou, M.; Sala, H. (GEAP)**.**;**Snower** , **D. J.**
"The macroeconomics of the labor market: Three fundamental views"
(Desembre 2006)

2007

XREAP2007-01
Castany, L (AQR-IREA)**; López-Bazo, E.** (AQR-IREA)**.**;**Moreno** , **R.** (AQR-IREA)
"Decomposing differences in total factor productivity across firm size"
(Març 2007)

XREAP2007-02
Raymond, J. Ll. (GEAP)**; Roig, J. Ll.** (GEAP)
"Una propuesta de evaluación de las externalidades de capital humano en la empresa"
(Abril 2007)

XREAP2007-03
Durán, J. M. (IEB)**; Esteller, A.** (IEB)
"An empirical analysis of wealth taxation: Equity vs. Tax compliance"
(Juny 2007)

XREAP2007-04
Matas, A. (GEAP)**; Raymond, J.Ll.** (GEAP)
"Cross-section data, disequilibrium situations and estimated coefficients: evidence from car ownership demand"
(Juny 2007)

XREAP2007-05
Jofre-Montseny, J. (IEB)**; Solé-Ollé, A.** (IEB)
"Tax differentials and agglomeration economies in intraregional firm location"
(Juny 2007)

XREAP2007-06
Álvarez-Albelo, C. (CREB)**; Hernández-Martín, R.**
"Explaining high economic growth in small tourism countries with a dynamic general equilibrium model"
(Juliol 2007)

XREAP2007-07
Duch, N. (IEB)**; Montolio, D.** (IEB); **Mediavilla, M.**
"Evaluating the impact of public subsidies on a firm's performance: a quasi-experimental approach"
(Juliol 2007)

XREAP2007-08
Segarra-Blasco, A. (GRIT)
"Innovation sources and productivity: a quantile regression analysis"
(Octubre 2007)

XREAP2007-09
Albalate, D. (PPRE-IREA)
"Shifting death to their Alternatives: The case of Toll Motorways"
(Octubre 2007)

XREAP2007-10
Segarra-Blasco, A. (GRIT); **Garcia-Quevedo, J.** (IEB); **Teruel-Carrizosa, M.** (GRIT)
"Barriers to innovation and public policy in catalonia"
(Novembre 2007)

XREAP2007-11
Bel, G. (PPRE-IREA); **Foote, J.**
"Comparison of recent toll road concession transactions in the United States and France"
(Novembre 2007)

XREAP2007-12
Segarra-Blasco, A. (GRIT);
"Innovation, R&D spillovers and productivity: the role of knowledge-intensive services"
(Novembre 2007)

XREAP2007-13
Bermúdez Morata, Ll. (RFA-IREA); **Guillén Estany, M.** (RFA-IREA), **Solé Auró, A**. (RFA-IREA)
"Impacto de la inmigración sobre la esperanza de vida en salud y en discapacidad de la población española"
(Novembre 2007)

XREAP2007-14
Calaeys, P. (AQR-IREA); **Ramos, R.** (AQR-IREA), **Suriñach, J**. (AQR-IREA)
"Fiscal sustainability across government tiers"
(Desembre 2007)

XREAP2007-15
Sánchez Hugalbe, A. (IEB)
"Influencia de la inmigración en la elección escolar"
(Desembre 2007)

2008

XREAP2008-01
Durán Weitkamp, C. (GRIT)**; Martín Bofarull, M.** (GRIT) ; **Pablo Martí, F.**
"Economic effects of road accessibility in the Pyrenees: User perspective"
(Gener 2008)

XREAP2008-02
Díaz-Serrano, L.; Stoyanova, A. P. (CREB)
"The Causal Relationship between Individual's Choice Behavior and Self-Reported Satisfaction: the Case of Residential Mobility in the EU"
(Març 2008)

XREAP2008-03
Matas, A. (GEAP)**; Raymond, J. L.** (GEAP)**; Roig, J. L.** (GEAP)
"Car ownership and access to jobs in Spain"
(Abril 2008)

XREAP2008-04
Bel, G. (PPRE-IREA) **; Fageda, X.** (PPRE-IREA)
"Privatization and competition in the delivery of local services: An empirical examination of the dual market hypothesis"
(Abril 2008)

XREAP2008-05
Matas, A. (GEAP); **Raymond, J. L.** (GEAP); **Roig, J. L.** (GEAP)
"**Job accessibility and employment probability**"
(Maig 2008)

XREAP2008-06
Basher, S. A.; Carrión, J. Ll. (AQR-IREA)
Deconstructing Shocks and Persistence in OECD Real Exchange Rates
(Juny 2008)

XREAP2008-07
Sanromá, E. (IEB)**; Ramos, R.** (AQR-IREA); Simón, H.
Portabilidad del capital humano y asimilación de los inmigrantes. Evidencia para España
(Juliol 2008)

XREAP2008-08
Basher, S. A.; Carrión, J. Ll. (AQR-IREA)
Price level convergence, purchasing power parity and multiple structural breaks: An application to US cities
(Juliol 2008)

XREAP2008-09
Bermúdez, Ll. (RFA-IREA)
A priori ratemaking using bivariate poisson regression models
(Juliol 2008)

XREAP2008-10
Solé-Ollé, A. (IEB), **Hortas Rico, M.** (IEB)
Does urban sprawl increase the costs of providing local public services? Evidence from Spanish municipalities
(Novembre 2008)

XREAP2008-11
Teruel-Carrizosa, M. (GRIT), **Segarra-Blasco, A.** (GRIT)
Immigration and Firm Growth: Evidence from Spanish cities
(Novembre 2008)

XREAP2008-12
Duch-Brown, N. (IEB), **García-Quevedo, J.** (IEB), **Montolio, D.** (IEB)
Assessing the assignation of public subsidies: Do the experts choose the most efficient R&D projects?
(Novembre 2008)

XREAP2008-13
Bilotkach, V., **Fageda, X.** (PPRE-IREA), **Flores-Fillol, R.**
Scheduled service versus personal transportation: the role of distance
(Desembre 2008)

XREAP2008-14
Albalate, D. (PPRE-IREA), **Gel, G.** (PPRE-IREA)
Tourism and urban transport: Holding demand pressure under supply constraints
(Desembre 2008)

2009

XREAP2009-01
Calonge, S. (CREB)**; Tejada, O.**
"A theoretical and practical study on linear reforms of dual taxes"
(Febrer 2009)

XREAP2009-02
Albalate, D. (PPRE-IREA)**; Fernández-Villadangos, L.** (PPRE-IREA)
"Exploring Determinants of Urban Motorcycle Accident Severity: The Case of Barcelona"
(Març 2009)

XREAP2009-03
Borrell, J. R. (PPRE-IREA)**; Fernández-Villadangos, L.** (PPRE-IREA)
"Assessing excess profits from different entry regulations"
(Abril 2009)

XREAP2009-04
Sanromá, E. (IEB)**; Ramos, R.** (AQR-IREA), **Simon, H.**
"Los salarios de los inmigrantes en el mercado de trabajo español. ¿Importa el origen del capital humano?"
(Abril 2009)

XREAP2009-05
Jiménez, J. L.; Perdiguero, J. (PPRE-IREA)
"(No)competition in the Spanish retailing gasoline market: a variance filter approach"
(Maig 2009)

XREAP2009-06
Álvarez-Albelo,C. D. (CREB)**, Manresa, A.** (CREB)**, Pigem-Vigo, M.** (CREB)
"International trade as the sole engine of growth for an economy"
(Juny 2009)

XREAP2009-07
Callejón, M. (PPRE-IREA)**, Ortún V, M.**
"The Black Box of Business Dynamics"
(Setembre 2009)

XREAP2009-08
Lucena, A. (CREB)
"The antecedents and innovation consequences of organizational search: empirical evidence for Spain"
(Octubre 2009)

XREAP2009-09
Domènech Campmajó, L. (PPRE-IREA)
"Competition between TV Platforms"
(Octubre 2009)

XREAP2009-10
Solé-Auró, A. (RFA-IREA),**Guillén, M.** (RFA-IREA)**, Crimmins, E. M.**
"Health care utilization among immigrants and native-born populations in 11 European countries. Results from the Survey of Health, Ageing and Retirement in Europe"
(Octubre 2009)

XREAP2009-11
Segarra, A. (GRIT)**, Teruel, M.** (GRIT)
"Small firms, growth and financial constraints"
(Octubre 2009)

XREAP2009-12
Matas, A. (GEAP)**, Raymond, J.Ll.** (GEAP), **Ruiz, A.** (GEAP)
"Traffic forecasts under uncertainty and capacity constraints"
(Novembre 2009)

XREAP2009-13
Sole-Ollé, A. (IEB)
"Inter-regional redistribution through infrastructure investment: tactical or programmatic?"
(Novembre 2009)

XREAP2009-14
Del Barrio-Castro, T., **García-Quevedo, J.** (IEB)
"The determinants of university patenting: Do incentives matter?"
(Novembre 2009)

XREAP2009-15
Ramos, R. (AQR-IREA), **Suriñach, J.** (AQR-IREA)**, Artís, M.** (AQR-IREA)
"Human capital spillovers, productivity and regional convergence in Spain"
(Novembre 2009)

XREAP2009-16
Álvarez-Albelo, C. D. (CREB), **Hernández-Martín, R.**
"The commons and anti-commons problems in the tourism economy"
(Desembre 2009)

2010

XREAP2010-01
García-López, M. A. (GEAP)
"The Accessibility City. When Transport Infrastructure Matters in Urban Spatial Structure"
(Febrer 2010)

XREAP2010-02
García-Quevedo, J. (IEB), **Mas-Verdú, F.** (IEB), **Polo-Otero, J.** (IEB)
"Which firms want PhDs? The effect of the university-industry relationship on the PhD labour market"
(Març 2010)

XREAP2010-03
Pitt, D., **Guillén, M.** (RFA-IREA)
"An introduction to parametric and non-parametric models for bivariate positive insurance claim severity distributions"
(Març 2010)

XREAP2010-04
Bermúdez, Ll. (RFA-IREA), **Karlis, D.**
"Modelling dependence in a ratemaking procedure with multivariate Poisson regression models"
(Abril 2010)

XREAP2010-05
Di Paolo, A. (IEB)
"Parental education and family characteristics: educational opportunities across cohorts in Italy and Spain"
(Maig 2010)

XREAP2010-06
Simón, H. (IEB), **Ramos, R.** (AQR-IREA), **Sanromá, E.** (IEB)
"Movilidad ocupacional de los inmigrantes en una economía de bajas cualificaciones. El caso de España"
(Juny 2010)

XREAP2010-07
Di Paolo, A. (GEAP & IEB), **Raymond, J. Ll.** (GEAP & IEB)
"Language knowledge and earnings in Catalonia"
(Juliol 2010)

XREAP2010-08
Bolancé, C. (RFA-IREA), **Alemany, R.** (RFA-IREA), **Guillén, M.** (RFA-IREA)
"Prediction of the economic cost of individual long-term care in the Spanish population"
(Setembre 2010)

XREAP2010-09
Di Paolo, A. (GEAP & IEB)
"Knowledge of catalan, public/private sector choice and earnings: Evidence from a double sample selection model"
(Setembre 2010)

XREAP2010-10
Coad, A., **Segarra, A.** (GRIT), **Teruel, M.** (GRIT)
"Like milk or wine: Does firm performance improve with age?"
(Setembre 2010)

XREAP2010-11
Di Paolo, A. (GEAP & IEB), **Raymond, J. Ll.** (GEAP & IEB), **Calero, J.** (IEB)
"Exploring educational mobility in Europe"
(Octubre 2010)

XREAP2010-12
Borrell, A. (GiM-IREA), **Fernández-Villadangos, L.** (GiM-IREA)
"Clustering or scattering: the underlying reason for regulating distance among retail outlets"
(Desembre 2010)

XREAP2010-13
Di Paolo, A. (GEAP & IEB)
"School composition effects in Spain"
(Desembre 2010)

XREAP2010-14
Fageda, X. (GiM-IREA), **Flores-Fillol, R.**
"Technology, Business Models and Network Structure in the Airline Industry"
(Desembre 2010)

XREAP2010-15
Albalate, D. (GiM-IREA), **Bel, G.** (GiM-IREA), **Fageda, X.** (GiM-IREA)
"Is it Redistribution or Centralization? On the Determinants of Government Investment in Infrastructure"
(Desembre 2010)

XREAP2010-16
Oppedisano, V., **Turati, G.**
"What are the causes of educational inequalities and of their evolution over time in Europe? Evidence from PISA"
(Desembre 2010)

XREAP2010-17
Canova, L., **Vaglio, A.**
"Why do educated mothers matter? A model of parental help"
(Desembre 2010)

2011

XREAP2011-01
Fageda, X. (GiM-IREA), **Perdiguero, J.** (GiM-IREA)
"An empirical analysis of a merger between a network and low-cost airlines"
(Maig 2011)

XREAP2011-02
Moreno-Torres, I. (ACCO, CRES & GiM-IREA)
"What if there was a stronger pharmaceutical price competition in Spain? When regulation has a similar effect to collusion"
(Maig 2011)

XREAP2011-03
Miguélez, E. (AQR-IREA); **Gómez-Miguélez, I.**
"Singling out individual inventors from patent data"
(Maig 2011)

XREAP2011-04
Moreno-Torres, I. (ACCO, CRES & GiM-IREA)
"Generic drugs in Spain: price competition vs. moral hazard"
(Maig 2011)

XREAP2011-05
Nieto, S. (AQR-IREA)**, Ramos, R.** (AQR-IREA)
"¿Afecta la sobreeducación de los padres al rendimiento académico de sus hijos?"
(Maig 2011)

XREAP2011-06
Pitt, D., **Guillén, M.** (RFA-IREA), **Bolancé, C.** (RFA-IREA)
"Estimation of Parametric and Nonparametric Models for Univariate Claim Severity Distributions - an approach using R"
(Juny 2011)

XREAP2011-07
Guillén, M. (RFA-IREA), **Comas-Herrera, A.**
"How much risk is mitigated by LTC Insurance? A case study of the public system in Spain"
(Juny 2011)

XREAP2011-08
Ayuso, M. (RFA-IREA), **Guillén, M.** (RFA-IREA), **Bolancé, C.** (RFA-IREA)
"Loss risk through fraud in car insurance"
(Juny 2011)

XREAP2011-09
Duch-Brown, N. (IEB), **García-Quevedo, J.** (IEB), **Montolio, D.** (IEB)
"The link between public support and private R&D effort: What is the optimal subsidy?"
(Juny 2011)

XREAP2011-10
Bermúdez, Ll. (RFA-IREA), **Karlis, D.**
"Mixture of bivariate Poisson regression models with an application to insurance"
(Juliol 2011)

XREAP2011-11
Varela-Irimia, X-L. (GRIT)
"Age effects, unobserved characteristics and hedonic price indexes: The Spanish car market in the 1990s"
(Agost 2011)

XREAP2011-12
Bermúdez, Ll. (RFA-IREA), **Ferri, A.** (RFA-IREA), **Guillén, M.** (RFA-IREA)
"A correlation sensitivity analysis of non-life underwriting risk in solvency capital requirement estimation"
(Setembre 2011)

XREAP2011-13
Guillén, M. (RFA-IREA), **Pérez-Marín, A.** (RFA-IREA), **Alcañiz, M.** (RFA-IREA)
"A logistic regression approach to estimating customer profit loss due to lapses in insurance"
(Octubre 2011)

XREAP2011-14
Jiménez, J. L., **Perdiguero, J.** (GiM-IREA), **García, C.**
"Evaluation of subsidies programs to sell green cars: Impact on prices, quantities and efficiency"
(Octubre 2011)

XREAP2011-15
Arespa, M. (CREB)
"A New Open Economy Macroeconomic Model with Endogenous Portfolio Diversification and Firms Entry"
(Octubre 2011)

XREAP2011-16
Matas, A. (GEAP), **Raymond, J. L.** (GEAP), **Roig, J.L.** (GEAP)
"The impact of agglomeration effects and accessibility on wages"
(Novembre 2011)

XREAP2011-17
Segarra, A. (GRIT)
"R&D cooperation between Spanish firms and scientific partners: what is the role of tertiary education?"
(Novembre 2011)

XREAP2011-18
García-Pérez, J. I.; **Hidalgo-Hidalgo, M.**; **Robles-Zurita, J. A.**
"Does grade retention affect achievement? Some evidence from PISA"
(Novembre 2011)

XREAP2011-19
Arespa, M. (CREB)
"Macroeconomics of extensive margins: a simple model"
(Novembre 2011)

XREAP2011-20
García-Quevedo, J. (IEB), **Pellegrino, G.** (IEB), **Vivarelli, M.**
"The determinants of YICs' R&D activity"
(Desembre 2011)

XREAP2011-21
González-Val, R. (IEB), **Olmo, J.**
"Growth in a Cross-Section of Cities: Location, Increasing Returns or Random Growth?"
(Desembre 2011)

XREAP2011-22
Gombau, V. (GRIT), **Segarra, A.** (GRIT)
"The Innovation and Imitation Dichotomy in Spanish firms: do absorptive capacity and the technological frontier matter?"
(Desembre 2011)

2012

XREAP2012-01
Borrell, J. R. (GiM-IREA), **Jiménez, J. L.**, **García, C.**
"Evaluating Antitrust Leniency Programs"
(Gener 2012)

XREAP2012-02
Ferri, A. (RFA-IREA), **Guillén, M.** (RFA-IREA), **Bermúdez, Ll.** (RFA-IREA)
"Solvency capital estimation and risk measures"
(Gener 2012)

XREAP2012-03
Ferri, A. (RFA-IREA), **Bermúdez, Ll.** (RFA-IREA), **Guillén, M.** (RFA-IREA)
"How to use the standard model with own data"
(Febrer 2012)

XREAP2012-04
Perdiguero, J. (GiM-IREA), **Borrell, J.R.** (GiM-IREA)
"Driving competition in local gasoline markets"
(Març 2012)

XREAP2012-05
D'Amico, G., **Guillen, M.** (RFA-IREA), Manca, R.
"Discrete time Non-homogeneous Semi-Markov Processes applied to Models for Disability Insurance"
(Març 2012)

XREAP2012-06
Bové-Sans, M. A. (GRIT), Laguado-Ramírez, R.
"Quantitative analysis of image factors in a cultural heritage tourist destination"
(Abril 2012)

XREAP2012-07
Tello, C. (AQR-IREA), **Ramos, R.** (AQR-IREA), **Artís, M.** (AQR-IREA)
"Changes in wage structure in Mexico going beyond the mean: An analysis of differences in distribution, 1987-2008"
(Maig 2012)

XREAP2012-08
Jofre-Monseny, J. (IEB), **Marín-López, R.** (IEB), **Viladecans-Marsal, E.** (IEB)
"What underlies localization and urbanization economies? Evidence from the location of new firms"
(Maig 2012)

XREAP2012-09
Muñiz, I. (GEAP), **Calatayud, D.**, **Dobaño, R.**
"Los límites de la compacidad urbana como instrumento a favor de la sostenibilidad. La hipótesis de la compensación en Barcelona medida a través de la huella ecológica de la movilidad y la vivienda"
(Maig 2012)

XREAP2012-10
Arqué-Castells, P. (GEAP), **Mohnen, P.**
"Sunk costs, extensive R&D subsidies and permanent inducement effects"
(Maig 2012)

XREAP2012-11
Boj, E. (CREB), **Delicado, P.**, **Fortiana, J**., **Esteve, A.**, **Caballé, A.**
"Local Distance-Based Generalized Linear Models using the dbstats package for R"
(Maig 2012)

XREAP2012-12
Royuela, V. (AQR-IREA)
"What about people in European Regional Science?"
(Maig 2012)

XREAP2012-13
Osorio A. M. (RFA-IREA)**, Bolancé, C.** (RFA-IREA), **Madise, N.**
"Intermediary and structural determinants of early childhood health in Colombia: exploring the role of communities"
(Juny 2012)

XREAP2012-14
Miguelez. E. (AQR-IREA)**, Moreno, R.** (AQR-IREA)
"Do labour mobility and networks foster geographical knowledge diffusion? The case of European regions"
(Juliol 2012)

XREAP2012-15
Teixidó-Figueras, J. (GRIT)**, Duró, J. A.** (GRIT)
"Ecological Footprint Inequality: A methodological review and some results"
(Setembre 2012)

XREAP2012-16
Varela-Irimia, X-L. (GRIT)
"Profitability, uncertainty and multi-product firm product proliferation: The Spanish car industry"
(Setembre 2012)

XREAP2012-17
Duró, J. A. (GRIT), **Teixidó-Figueras, J.** (GRIT)
"Ecological Footprint Inequality across countries: the role of environment intensity, income and interaction effects"
(Octubre 2012)

XREAP2012-18
Manresa, A. (CREB), **Sancho, F.**
"Leontief versus Ghosh: two faces of the same coin"
(Octubre 2012)

XREAP2012-19
Alemany, R. (RFA-IREA), **Bolancé, C.** (RFA-IREA), **Guillén, M.** (RFA-IREA)
"Nonparametric estimation of Value-at-Risk"
(Octubre 2012)

XREAP2012-20
Herrera-Idárraga, P. (AQR-IREA), **López-Bazo, E.** (AQR-IREA), **Motellón, E.** (AQR-IREA)
"Informality and overeducation in the labor market of a developing country"
(Novembre 2012)

XREAP2012-21
Di Paolo, A. (AQR-IREA)
"(Endogenous) occupational choices and job satisfaction among recent PhD recipients: evidence from Catalonia"
(Desembre 2012)

2013

XREAP2013-01
Segarra, A. (GRIT), **García-Quevedo, J.** (IEB), **Teruel, M.** (GRIT)
"Financial constraints and the failure of innovation projects"
(Març 2013)

XREAP2013-02
Osorio, A. M. (RFA-IREA), **Bolancé, C.** (RFA-IREA), Madise, N., Rathmann, K.
"Social Determinants of Child Health in Colombia: Can Community Education Moderate the Effect of Family Characteristics?"
(Març 2013)

XREAP2013-03
Teixidó-Figueras, J. (GRIT), **Duró, J. A.** (GRIT)
"The building blocks of international ecological footprint inequality: a regression-based decomposition"
(Abril 2013)

XREAP2013-04
Salcedo-Sanz, S., **Carro-Calvo, L.**, **Claramunt, M.** (CREB), **Castañer, A.** (CREB), **Marmol, M.** (CREB)
"An Analysis of Black-box Optimization Problems in Reinsurance: Evolutionary-based Approaches"
(Maig 2013)

XREAP2013-05
Alcañiz, M. (RFA), **Guillén, M.** (RFA), **Sánchez-Moscona, D.** (RFA), **Santolino, M.** (RFA), **Llatje, O., Ramon, Ll.**
"Prevalence of alcohol-impaired drivers based on random breath tests in a roadside survey"
(Juliol 2013)

XREAP2013-06
Matas, A. (GEAP & IEB), **Raymond, J. Ll.** (GEAP & IEB), **Roig, J. L.** (GEAP)
"How market access shapes human capital investment in a peripheral country"
(Octubre 2013)

XREAP2013-07
Di Paolo, A. (AQR-IREA), **Tansel, A.**
"Returns to Foreign Language Skills in a Developing Country: The Case of Turkey"
(Novembre 2013)

XREAP2013-08
Fernández Gual, V. (GRIT), **Segarra, A.** (GRIT)
"The Impact of Cooperation on R&D, Innovation andProductivity: an Analysis of Spanish Manufacturing and Services Firms"
(Novembre 2013)

XREAP2013-09
Bahraoui, Z. (RFA); **Bolancé, C.** (RFA); **Pérez-Marín. A. M.** (RFA)
"Testing extreme value copulas to estimate the quantile"
(Novembre 2013)

2014

XREAP2014-01
Solé-Auró, A. (RFA), **Alcañiz, M.** (RFA)
"Are we living longer but less healthy? Trends in mortality and morbidity in Catalonia (Spain), 1994-2011"
(Gener 2014)

XREAP2014-02
Teixidó-Figueres, J. (GRIT), **Duro, J. A.** (GRIT)
"Spatial Polarization of the Ecological Footprint distribution"
(Febrer 2014)

XREAP2014-03
Cristobal-Cebolla, A.; Gil Lafuente, A. M. (RFA), **Merigó Lindhal, J. M.** (RFA)
"La importancia del control de los costes de la no-calidad en la empresa"
(Febrer 2014)

XREAP2014-04
Castañer, A. (CREB); **Claramunt, M.M.** (CREB)
"Optimal stop-loss reinsurance: a dependence analysis"
(Abril 2014)

XREAP2014-05
Di Paolo, A. (AQR-IREA); **Matas, A.** (GEAP); **Raymond, J. Ll.** (GEAP)
"Job accessibility, employment and job-education mismatch in the metropolitan area of Barcelona"
(Maig 2014)

XREAP2014-06
Di Paolo, A. (AQR-IREA); **Mañé, F.**
"Are we wasting our talent? Overqualification and overskilling among PhD graduates"
(Juny 2014)

XREAP2014-07
Segarra, A. (GRIT); **Teruel, M.** (GRIT); **Bové, M. A.** (GRIT)
"A territorial approach to R&D subsidies: Empirical evidence for Catalonian firms"
(Setembre 2014)

XREAP2014-08
Ramos, R. (AQR-IREA); **Sanromá, E.** (IEB); **Simón, H.**
"Public-private sector wage differentials by type of contract: evidence from Spain"
(Octubre 2014)

XREAP2014-09
Bel, G. (GiM-IREA); **Bolancé, C.** (Riskcenter-IREA); **Guillén, M.** (Riskcenter-IREA)**; Rosell, J.** (GiM-IREA)
"The environmental effects of changing speed limits: a quantile regression approach"
(Desembre 2014)

2015

XREAP2015-01
Bolance, C. (Riskcenter-IREA); **Bahraoui, Z.** (Riskcenter-IREA), **Alemany, R.** (Risckcenter-IREA)
"Estimating extreme value cumulative distribution functions using bias-corrected kernel approaches"
(Gener 2015)

XARXA DE REFERÈNCIA EN ECONOMIA APLICADA

xarxa.xreap@gmail.com