DOCUMENT DE TREBALL

XREAP2012-19

Nonparametric estimation of Value-at-Risk

Ramon Alemany (RFA-IREA, XREAP)
Catalina Bolancé (RFA-IREA, XREAP)
Montserrat Guillén (RFA-IREA, XREAP)

XARXA DE REFERÈNCIA EN ECONOMIA APLICADA

Nonparametric estimation of Value-at-Risk

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

October 16, 2012

Abstract

A method to estimate an extreme quantile that requires no distributional assumptions is presented. The approach is based on transformed kernel estimation of the cumulative distribution function (cdf). The proposed method consists of a double transformation kernel estimation. We derive optimal bandwidth selection methods that have a direct expression for the smoothing parameter. The bandwidth can accommodate to the given quantile level. The procedure is useful for large data sets and improves quantile estimation compared to other methods in heavy tailed distributions. Implementation is straightforward and R programs are available.

Keywords: kernel estimation, bandwidth selection, quantile, risk measures.

1 Introduction

Risk measures and their mathematical properties have been widely studied in the literature (see, for instance, the books by McNeil et al. (2005) and Jorion (2007) or articles such as Dhaene et al. (2006) among many others). Most of those contributions and applications in risk management usually assume a parametric distribution for the loss random variable[1]. Deviations from parametric hypothesis can be critical in the extremes and produce inaccurate results (see, Kupiec, 1995). Krätschmer and Zähle (2011) investigated the error made even when the normal approximation is plugged in a general distribution-invariant risk measure.

[1]Standard industry models as CreditRisk$^+$ are parametric. See, Fan and Gu (2003), and references therein for semiparametric models.

Our approach is nonparametric as in Peng et al. (2012), Cai and Wang (2008) and Jones and Zitikis (2007). We propose a method to estimate quantiles that is based on a nonparametric estimate of the cumulative distribution function with an optimal bandwidth at the desired quantile level. Eling (2012) recently used a similar benchmark nonparametric fit to describe claims severity distributions in property-liability insurance (see, Bolancé et al., 2012b, for details) but the choice of the smoothing parameter needs further analysis. Besides, Eling (2012) was interested in the fit for the density of claims severity, not on risk measurement or quantiles[2]. We present the nonparametric estimation approach and focus on the bandwidth choice. We also carry out a simulation exercise.

A risk measure widely used to quantify the risk is the value-at-risk with level α. It is defined as follows,

$$VaR_{\alpha}(X) = \inf\{x, F_X(x) \geq \alpha\} = F_X^{-1}(\alpha), \quad (1)$$

where X is a random variable with probability distribution function (pdf) f_X, and cumulative distribution function (cdf) F_X. Artzner et al. (1999) discussed other risk measures, but they stated that expected shortfall is preferred in practice due to its better properties, although value-at-risk is widely used in applications.

The VaR_{α} is used both as an internal risk management tool and as a regulatory measure of risk exposure to calculate capital adequacy requirements in financial and insurance institutions. In this paper we propose a method to estimate the VaR_{α} in extreme quantiles, based on transformed kernel estimation (TKE) of the cdf of losses. The proposed method consists of a double transformation kernel estimation (DTKE), and it works well for very extreme levels and a large sample size. It also improves quantile estimation compared to existing methods. An additional contribution is that we propose a simple expression for an optimal bandwidth parameter. Thus, we advocate that there is little advantage of assuming parametric distributions when calculating value-at-risk for heavy tailed data, given that the nonparametric approach implementation is very straightforward.

Some previous research has already studied nonparametric estimation of quantiles. On the one hand Azzalini (1981) suggested to estimate the cdf and then to obtain the quantile from its inverse function. On the other hand Harrell and Davis (1982) proposed an alternative quantile estimator, based in a weighted sum of sample observations. Later, Sheather and Marron (1990) analysed the existing kernel methods for quantile estimation and proposed a smoothing parameter. None of those contributions, however, focused on highly skewed or heavy tailed distribution, which most often appear in financial and insurance risk management.

[2]Eling (2012) worked with two empirical data sets. The first dataset is US indemnity losses and the second is comprised of Danish fire losses. His work indicated that the transformation kernel (Bolancé et al., 2003) is the best and second best approach when compared with the parametric distributions in terms of the log likelihood value in his applications. The transformation kernel approach performed extremely well there and confirmed the results presented by Bolancé et al. (2008a) for auto insurance.

Recently, Swanepoel and Van Graan (2005) presented kernel estimation of a cdf using nonparametric transformation, i.e. a simple form of transformed kernel estimation. Instead, Bolancé et al. (2008b) used a parametric transformation, which provides good results in the estimation of conditional tail expectation. Here, we propose an improved nonparametric procedure to estimate the VaR_α in finance and insurance applications and derive an optimal expression for the bandwidth parameter.

2 Motivation and outline

Our motivation is found on the statistical assumptions underlying the random behaviour of loss distributions. In practice, calculating VaR_α requires to assume a particular stochastic behaviour of losses. Assumptions have generally been based on three possible statistical principles: i) the empirical statistical distribution of the loss or some smoothed version, ii) assuming that the loss follows a Normal or Student t distribution and iii) some other alternative parametric approximations. Sample size is a key factor to determine the method to estimate the quantile. In order to use the empirical distribution function, a minimum sample size is required. The Normal approximation provides a straightforward expression for the most popular risk measures, although the loss may be far from having a Normal shape or even a Student t distribution. Alternatively, one should find a suitable heavy tailed parametric distribution to which the loss data should fit (see, for example, McNeil et al., 2005; Jorion, 2007; Bolancé et al., 2012b). Extreme value theory can be used to locate the tail of the distribution (see, Reiss and Thomas, 1997; Hill, 1975; Guillén et al., 2011).

A principal difference between our transformed kernel estimation and the fit of a heavy tailed parametric loss distribution is that we use sample information to estimate the parameters of a initial parametric model and, later, we also use the sample information to correct this initial fit. The proposed method works when losses have heavy tailed distributions, it is easy to implement and it provides consistent results. It is very flexible, so it is comparable to the empirical distribution approach. We can affirm that the method proposed in this work smooths the shape of the empirical distribution and extrapolates its behaviour when dealing with extremes, where data are very scarce or non existent.

The results of our simulation study show that our double transformed kernel estimation method can be applied to risk measurement and is specially suitable when the sample size is large. This is useful when basic parametric densities provide a poor fit in the tail. In the transformed kernel approach, no parametric form is imposed on the loss distribution, but, most importantly, this method avoids defining where the tail of the loss distribution starts in order to apply extreme value theory.

When writing this article, we decided to summarize basic nonparametric concepts that appear quite frequently elsewhere[3]. We introduce kernel estimation notation to make the presentation self-contained. In Section 3 we present nonparametric estimation of a pdf and a cdf. We also describe nonparametric estimation of cdf in connection with estimation of value-at-risk. Section 4 introduces transformation kernel estimation of a cdf and a new result on its asymptotic properties. Double transformation kernel estimation of a cdf and the selection of the smoothing parameter are studied in Section 5. Section 6 presents a simulation study where we can confirm the properties of the methods proposed in the previous sections. The most relevant conclusions and a discussion are given in the last section. Implementation tools in R are available from the authors and detailed hands-on examples of transformation kernel estimation can be found in Bolancé et al. (2012b).

3 Nonparametric estimation of a cumulative distribution function

Let X be a random variable which represents a loss amount; its cdf is F_X. Let us assume that X_i $i = 1, ..., n$ denotes data observations from the loss random variable X. For instance, loss data may also arise from historical simulation or they may have been generated in a Monte Carlo analysis. A natural nonparametric method to estimate cdf is the empirical distribution,

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x), \tag{2}$$

where $I(\cdot) = 1$ if condition between parentheses is true. Then, the empirical estimate of value-at-risk is:

$$VaR_\alpha(X) = \inf\left\{x, \widehat{F}_n(x) \geq \alpha\right\}. \tag{3}$$

Estimation of the empirical distribution is very simple, but it cannot extrapolate beyond the maximum observed data point. This is especially troublesome if the sample is not too large, and one may suspect that the probability of a loss larger than the maximum observed loss in the data sample is not zero.

Classical kernel estimation (CKE) of cdf F_X is obtained by integration of the classical kernel estimation of its pdf f_X. By means of a change of variable, the usual expression for the kernel

[3] Many recent contributions in insurance are based on nonparametric statistical methods. For instance, Lopez (2012) provided a new nonparametric estimator of the joint distribution of two lifetimes for mortality analysis and Kim (2010) studied the bias of the empirical distorsion risk measure estimate.

estimator of a cdf is obtained:

$$\begin{aligned} \widehat{F}_X(x) &= \int_{-\infty}^{x} \widehat{f}_X(u)du = \int_{-\infty}^{x} \frac{1}{nb} \sum_{i=1}^{n} k\left(\frac{u-X_i}{b}\right) du \\ &= \frac{1}{n} \sum_{i=1}^{n} \int_{-\infty}^{\frac{x-X_i}{b}} k(t)\, dt = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x-X_i}{b}\right), \end{aligned} \tag{4}$$

where $k(\cdot)$ is a pdf, which is known as the kernel function. It is usually a symmetric pdf, but this does not imply that the final estimate of F_X is symmetric. Some examples of very common kernel functions are the Epanechnikov and the Gaussian kernel (see, Silverman, 1986). Parameter b is the *bandwidth* or the smoothing parameter. It controls the smoothness of the cdf estimate. The larger b is, the smoother the resulting cdf. Function $K(\cdot)$ is the cdf of $k(\cdot)$.

The classical kernel estimation of a cdf as defined in (4) is not much different to the expression of the well-known empirical distribution in (2). Indeed, in (4) one should replace $K\left(\frac{x-X_i}{b}\right)$ by $I(X_i \leq x)$ in order to obtain (2). The main difference between (2) and (4) is that the empirical cdf only uses data below x to obtain the point estimate of $F_X(x)$, while the classical kernel cdf estimator uses all the data above and below x. In other words, the empirical cdf gives more weight to the observations that are smaller than x than it does to the observations that are larger than x.

In practice, to estimate VaR_α from $\widehat{F}_X(\cdot)$, we use the Newton-Raphson method to solve the equation:

$$\widehat{F}_X(x) = \alpha. \tag{5}$$

Properties of kernel cdf estimator were analyzed by Reiss (1981) and Azzalini (1981). Both point out that when $n \to \infty$, the mean squared error (MSE) of $\widehat{F}_X(x)$ can be approximated by:

$$\begin{aligned} E\left\{\widehat{F}_X(x) - F_X(x)\right\}^2 &\sim \frac{F_X(x)[1-F_X(x)]}{n} - f_X(x)\frac{b}{n}\left(1 - \int_{-1}^{1} K^2(t)\, dt\right) \\ &\quad + b^4\left(\frac{1}{2} f_X'(x) \int t^2 k(t)\, dt\right)^2 \\ &= \frac{F_X(x)[1-F_X(x)]}{n} - u(x) + b^4 v(x), \end{aligned} \tag{6}$$

where as in Azzalini (1981)

$$u(x) = f_X(x)\frac{b}{n}\left(1 - \int_{-1}^{1} K^2(t)\, dt\right)$$

and

$$v(x) = \left(\frac{1}{2} f_X'(x) \int t^2 k(t)\, dt\right)^2.$$

Expression (6) comes from a Taylor expansion of $\widehat{F}_X(x)$. The first two terms in (6) correspond to the asymptotic variance and the third term is the squared asymptotic bias. If (6) is compared to the MSE of the empirical distribution, which equals $(F_X(x)[1-F_X(x)])/n$, we conclude that the kernel cdf estimator has less variance than the empirical distribution estimator, but it has some bias which tends to zero if the sample size is large. Azzalini (1981) showed that the properties of CKE of cdf are transferred to the quantile estimator.

The value for the smoothing parameter b that minimizes (6) is:

$$b_x^* = \left(\frac{f_X(x) \int K(t)[1-K(t)]\, dt}{\left(f'_X(x) \int t^2 k(t)\, dt\right)^2} \right)^{\frac{1}{3}} n^{-\frac{1}{3}}. \tag{7}$$

Azzalini (1981) showed that (7) is also optimal when calculating the quantiles by solving (5). However, in practice, calculating b_x^* is not simple because it depends on the true value of $f_X(x)$. It is common to replace the theoretical value of $f_X(x)$ by the Normal pdf with zero mean and scale parameter σ, which is estimated from the raw data. However, as our goal is to estimate the quantile (i.e. the point x where the value of the cdf is $F(x)=\alpha$), this particular approach to estimating b_x^* seems unstable, as it depends on how precise is the estimate of f_X in the tail.

An alternative to the smoothing parameter defined in (7) is to use a value for the bandwidth that is asymptotically optimal for the entire domain of the cdf. So, we can estimate the optimal smoothing parameter for the entire domain. One way is to adapt the method described in Silverman (1986). We emphasize that this method starts from minimizing the mean integrated squared error (MISE):

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

The asymptotic value of MISE is known as A-MISE (asymptotic mean integrated squared error). When integrating the asymptotic expression of mean squared error given in (6), it follows that A-MISE is:

$$\frac{1}{n}\int F_X(x)[1-F_X(x)]\, dx - \frac{1}{n} b \int K(t)[1-K(t)]\, dt + \frac{1}{4} b^4 \int [f'_X(x)]^2 dx \left(\int t^2 k(t)\, dt\right)^2. \tag{8}$$

Minimizing (8) with respect to b, we find that the smoothing parameter which is asymptotically optimal for all the domain of the cdf is:

$$b^* = \left(\frac{\int K(t)[1-K(t)]\, dt}{\int [f'_X(x)]^2 dx \left(\int t^2 k(t)\, dt\right)^2} \right)^{\frac{1}{3}} n^{-\frac{1}{3}}. \tag{9}$$

Silverman (1986)) suggests to approximate (9) and replace the terms that depend on the theoretical density function by the value obtained when assuming that $\int [f'_X(x)]^2 dx$ can be estimated assuming a Normal distribution with parameters (μ, σ). Using the Epanechnikov kernel, the so called *rule-of-thumb bandwidth* is defined as:

$$\hat{b} = \sigma \left(\frac{180\sqrt{\pi}}{7} \right)^{\frac{1}{3}} n^{\frac{1}{3}} = 3.572 \sigma_X n^{-\frac{1}{3}}, \tag{10}$$

where, in practice, we replace σ_X by a consistent estimator from the available loss data.

Since the objective of this paper is to estimate a quantile in the tail of the distribution, as an alternative to (10) we can calculate the optimal smoothing parameter, so that more weight is given to the accuracy of the estimate in the part of the domain near the quantile. We assume that we estimate VaR_α for a level α close to 1, and we analyze the possibility of using the smoothing parameter based on the minimization of a weighted mean integrated squared error[4].

Proposition 1 *Let*

$$WISE\left\{\widehat{F}_X(x)\right\} = E\left\{\int \left[F_X(x) - \widehat{F}_X(x)\right]^2 x^2 dx\right\}.$$

Similarly to A-MISE, we denote as A-WISE the asymptotic value of WISE, that is equal to:

$$\frac{\int F_X(x)\left[1 - F_X(x)\right] x^2 dx}{n} - \frac{b \int f_X(x) x^2 dx \int K(t)\left[1 - K(t)\right] dt}{n} + \frac{1}{4} b^4 \int \left[f'_X(x)\right]^2 x^2 dx \left(\int t^2 k(t)\, dt\right)^2.$$

Minimizing the above expression with respect to the smoothing parameter b we find that the value of the bandwidth that minimizes A-WISE is equal to:

$$b^{**} = \left(\frac{\int f_X(x) x^2 dx \int K(t)\left[1 - K(t)\right] dt}{\int \left[f'_X(x)\right]^2 x^2 dx \left(\int t^2 k(t)\, dt\right)^2} \right)^{\frac{1}{3}} n^{\frac{1}{3}}. \tag{11}$$

Proposition 2 *In order to obtain a rule-of-thumb approximation of (11), we replace functionals $\int f_X(x) x^2 dx$ and $\int \left[f'_X(x)\right]^2 x^2 dx$ by their corresponding values if we assume that f_X is a Normal pdf with scale parameter σ and $k(\cdot)$ is the Epanechnikov kernel:*

$$\hat{b}^{**} = \sigma^{\frac{5}{3}} \left(\frac{8}{3}\right)^{\frac{1}{3}} n^{-\frac{5}{3}}. \tag{12}$$

[4]We assume that we are interested in the right tail, but a similar approach could be used for the left tail changing signs. Usually we work with non-negative losses, which is the typical setting in risk management of insurance claims.

Many authors have addressed the bandwidth selection problem. Sarda (1993) and Bowman et al. (1998) analyzed the choice of the smoothing parameter based on minimizing cross-validation function. All these methods are very much time-consuming and require rather lengthy implementation processes. That is the reason why we believe that none of them has been successful in practice. Altman and Léger (1995) discussed a plug-in method based on an expression of A-MISE which includes an additional term in the Taylor expansion of $\widehat{F}_X(x)$, but this leads to the need to estimate more than one functional expression and the results are still not straightforward. We will return to this problem in the next section, where a double transformation will guide the bandwidth choice.

As an alternative to kernel estimation of cdf, the kernel quantile estimator (KQE) is a classical method to estimate the VaR_α. Sheather and Marron (1990) reviewed different forms of obtaining KQE and they showed that they were all asymptotically equivalent. In order to compare the results with those obtained from the inverse of the kernel estimation of the distribution function, we also included Sheather and Marron (1990) approach in our simulation study. The kernel quantile estimator is:

$$KQ(\alpha) = \frac{\frac{1}{nb}\sum_{i=1}^{n} K\left(\frac{\frac{i-\frac{1}{2}}{n}-\alpha}{b}\right) X_{(i)}}{\frac{1}{nb}\sum_{i=1}^{n} K\left(\frac{\frac{i-\frac{1}{2}}{n}-\alpha}{b}\right)}. \tag{13}$$

In (13) we could use the bandwidth proposed in Harrell and Davis (1982) using a Gaussian kernel. So, the bandwidth would be:

$$b = \left[\frac{\alpha(1-\alpha)}{n+1}\right]^{\frac{1}{2}}. \tag{14}$$

Sheather and Marron (1990) proposed an optimal smoothing parameter for (13), but it involves great difficulty in calculations. These authors conducted a simulation exercise and compared the MSE of (13) using the bandwidth they proposed and the bandwidth given by (14). For quantiles close to 1, their results showed that the difference between the MSE of both proposals is not too large. So, we later use the expression given in (14).

4 Transformed Kernel Estimation

Let $T(\cdot)$ be a concave transformation where $Y = T(X)$ and $Y_i = T(X_i)$, $i = 1 \ldots n$ are the transformed observed losses. Then the kernel estimator of the transformed cumulative distribution function is:

$$\widehat{F}_Y(y) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{y-Y_i}{b}\right) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{T(x)-T(X_i)}{b}\right), \tag{15}$$

The transformated kernel estimation (TKE) of $F_X(x)$ is:

$$\widehat{F}_X(x) = \widehat{F}_{T(X)}(T(x)).$$

In expression (15), b and $K(\cdot)$ have already been defined in the previous section on the classical kernel estimation of the distribution function. In order to calculate VaR_α we use the Newton-Raphson method to solve the equation $\widehat{F}_{T(X)}(T(x)) = \alpha$ and once the result is obtained, we apply the inverse of the transformation.

To obtain the transformed kernel estimate, it is necessary to determine what transformation to use. Several authors have analyzed the transformation kernel estimation of the density function (see, Wand et al., 1991; Bolancé et al., 2003; Buch-Larsen et al., 2005; Pitt et al., 2012; Ruppert and Cline, 1994)[5]. However, few studies analyzed the transformed kernel estimate of the distribution function and the quantile (see, Swanepoel and Van Graan, 2005). In general, transformations are classified into parametric and nonparametric and, in turn, they may or may not correspond to a distribution function. The core objective of the transformation is that the chosen distribution of the new variable can be estimated using the classical kernel.

The work in Buch-Larsen et al. (2005) proposed to transform the data with the cdf associated with generalized Champernowne distribution. This is suitable for positive losses:

$$T(x) = \frac{(x+c)^\delta - c^\delta}{(x+c)^\delta + (M+c)^\delta - 2c^\delta}. \tag{16}$$

If we analyze the properties of this distribution, we can conclude that it has a very flexible shape. It is similar to a Lognormal in the low values and it tends to a Generalized Pareto in the extreme values. The estimation of transformation parameters is performed using the maximum likelihood method described in Buch-Larsen et al. (2005).

The transformed variable from a cdf follows a Uniform$(0,1)$ distribution. We know that when the value of density is larger than 0 at the boundary, classical kernel estimation of pdf does not integrate to 1. Thus, transformed kernel estimation of cdf cannot be used to estimate VaR_α when α is close to 1. To allow that the estimated pdf integrates to 1, Buch-Larsen et al.

[5]There are also some applied contributions in this area too (Guillen et al., 2007; Pinquet et al., 2001; Bolancé et al., 2008b; Bolancé et al., 2008, 2010a; Buch-Kromann et al., 2011; Englund et al., 2008; Bolancé et al., 2010b)

(2005) proposed using a boundary correction, that is:

$$\int\limits_{\max(-1,-y/b)}^{\min(1,(1-y)/b)} K(u)du. \tag{17}$$

Swanepoel and Van Graan (2005) proposed a transformed kernel estimation using a non-parametric cdf transformation:

$$T\left(x\right)=\frac{1}{n}\sum_{i=1}^{n}K\left(\frac{x-X_i}{b_0}\right), \tag{18}$$

where b_0 is the smoothing parameter for the nonparametric transformation. They also proposed to use the same bandwidth in the transformation and in the final estimation in (15). Similarly to (16), in (18) we also need to use the boundary correction in (17) to be able to calculate VaR_α when α is close to 1.

In Theorem 1 we analyze the effect of the transformation on the MSE of the transformation kernel estimation of $F_X(x)$ and so, on the value of the estimated quantile.

Theorem 1 *The MSE of the transformed kernel estimation of a cdf based on (15) is asymptotically equal to:*

$$\begin{aligned} &E\left\{\hat{F}_{T(X)}\left(T\left(x\right)\right)-F_{T(X)}\left(T\left(x\right)\right)\right\}^2 \\ &\quad= E_T\left\{\hat{F}_X\left(x\right)-F_X\left(x\right)\right\}^2 \\ &\sim \tfrac{F_X(x)[1-F_X(x)]}{n}-\tfrac{1}{T'(x)}f_X\left(x\right)\tfrac{b}{n}\left(1-\int_{-1}^{1}K^2\left(t\right)dt\right) \\ &\quad+\tfrac{1}{T'(x)}\left(1-\frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}}\right)^2\left[\tfrac{1}{2}f'_X\left(x\right)\int_{-1}^{1}t^2k\left(t\right)dt\right]^2 b^4 \\ &= \tfrac{F_X(x)[1-F_X(x)]}{n}-\tfrac{1}{T'(x)}u\left(x\right)+\tfrac{1}{T'(x)}\left(1-\frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}}\right)^2 v\left(x\right)b^4, \end{aligned} \tag{19}$$

where u and v are the same as before (see, Azzalini, 1981).

Proof 1 *Proof of Theorem 1 is in the Appendix.*

The result in Theorem 1 shows that the two last terms in MSE of TKE are a weighted sum of the two last terms in MSE of CKE.

Minimizing expression (19) with respect to b we obtain:

$$\begin{aligned} b_x^T &= T'(x)^{\frac{1}{3}} \left(1 - \frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}} \right)^{-\frac{2}{3}} \left(\frac{u(x)}{4v(x)} \right)^{\frac{1}{3}} n^{-\frac{1}{3}} \\ &= T'(x)^{\frac{1}{3}} \left(1 - \frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}} \right)^{-\frac{2}{3}} b_x^{Clas} = b_{T(x)}^{Clas}, \end{aligned}$$

where b_x^{Clas} is the optimal bandwidth in classical kernel estimation of $F_x(x)$ and $b_{T(x)}^{Clas}$ is the same for $F_{T(x)}(T(x))$. Replacing b_x^T in (19) we obtain the asymptotic optimal MSE:

$$\frac{F_X(x)\left[1 - F_X(x)\right]}{n} - T'(x)^{\frac{2}{3}} \left(1 - \frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}} \right)^{-\frac{2}{3}} \frac{5}{v(x)^{\frac{1}{3}}} \left(\frac{u(x)}{4n} \right)^{\frac{4}{3}}. \tag{20}$$

For the classical kernel estimation of $F_X(x)$, Azzalini (1981) finds that the asymptotic optimal MSE is:

$$\frac{F_X(x)\left[1 - F_X(x)\right]}{n} - \frac{5}{v(x)^{\frac{1}{3}}} \left(\frac{u(x)}{4n} \right)^{\frac{4}{3}}.$$

Then the MSE of transformed kernel estimation in (20) is smaller than MSE of classical estimation if:

$$T'(x)^{\frac{2}{3}} \left(1 - \frac{\frac{f_X(x)}{f'_X(x)}}{\frac{T'(x)}{T''(x)}} \right)^{-\frac{2}{3}} > 1. \tag{21}$$

In practice, the use of a suitable transformation reduces the variance at the expense of increasing the bias of the estimation. In the simulation presented later, we analyze to what extent this correction implies a reduction in mean square error of the transformed kernel estimate compared to the classical.

The optimal smoothing parameter to estimate the cdf using transformed kernel estimator coincides with the optimal smoothing parameter for the classical approach on the transformed variable. We can estimate the optimal smoothing parameter for the entire domain of the function using (10) or (11) replacing σ_X by σ_Y.

Moreover, we can approximate the optimal smoothing parameter in the quantile estimate by replacing the X by Y in the expression of b_x^* as defined in (7) and, conversely, $f_Y(y)$ by the value of the Normal density at point y. However, this approach provides a worse outcome than those expressed in (10) and (12) as the value of the Normal density at y is usually not an accurate approximation of the true density at that point, in heavy tailed distributions and high quantiles.

5 Double Transformed Kernel Estimation

The estimation that we describe in this section is based on the method proposed by Bolancé et al. (2008a) in the context of density functions. Here the objective is to estimate the cdf and the method is much simpler. We will also derive the corresponding quantile estimator.

In the expression of A-MISE given in (8), we can see that in order to obtain a smoothing parameter that is asymptotically optimal, it is sufficient to minimize:

$$\frac{1}{4}b^4 \int \left[f_Y'(y)\right]^2 dy \left(\int t^2 k(t)\, dt\right)^2 - \frac{1}{n} b \int K(t)\left[1 - K(t)\right] dt,$$

where, given b and $k(\cdot)$, the value is minimum when functional $\int \left[f_Y'(y)\right]^2 dy$ is minimum. Therefore, the proposed method is based on the transformation of the variable in order to achieve a distribution that minimizes the previous expression.

Terrell (1990) showed that the density of a $Beta\,(3,3)$ distribution defined on the domain $[-1,1]$ minimizes $\int \left[f_Y'(y)\right]^2 dy$, in the set of all densities with known variance. Its pdf and cdf are, respectively:

$$\begin{aligned} m(x) &= \frac{15}{16}\left(1 - x^2\right)^2, -1 \leq x \leq 1, \\ M(x) &= \frac{3}{16}x^5 - \frac{5}{8}x^3 + \frac{15}{16}x + \frac{1}{2}. \end{aligned} \tag{22}$$

The double transformation kernel estimation method requires an initial transformation of the data $T(X_i) = Z_i$, where we get a transformed variable distribution that is close to a Uniform$(0,1)$. Afterwards, the data are transformed again using the inverse of the distribution function of a $Beta\,(3,3)$, $M^{-1}(Z_i) = Y_i$. The resulting variable once the double transformation has been made, has a distribution that is close to a $Beta\,(3,3)$ (see, Bolancé, 2010). The double transformation kernel estimator (DTKE) is:

$$\begin{aligned} \widehat{F}_X(x) &= \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{M^{-1}(T(x)) - M^{-1}(T(X_i))}{b}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{y - Y_i}{b}\right). \end{aligned} \tag{23}$$

The smoothing parameter b can be calculated from expression (9) knowing that for the $Beta\,(3,3)$ distribution in (22) $\int \left[f_Y'\,(y)\right]^2 dy = 15/7$. So, using an Epanechnikov kernel, we obtain a simple smoothing parameter as follows:

$$\hat{b}^{*} = 3^{\frac{1}{3}} n^{-\frac{1}{3}}. \tag{24}$$

We can also compute the smoothing parameter using expression (11). If we substitute again $f_Y\,(y)$ by density $m\,(y)$ from a $Beta\,(3,3)$ as defined in (22), we obtain:

$$\hat{b}^{**} = (9/7)^{\frac{1}{3}}\, n^{-\frac{1}{3}}. \tag{25}$$

However, the distribution of the transformed variables has been established, with the pdf and the cdf defined in (22). It is crucial here, that this method provides an accurate way to obtain an estimation of the optimal value of the smoothing parameter precisely at the point where we wish to estimate the VaR_α. From expression (7), it follows that:

$$b_{T(x)}^{Clas} = \left(\frac{u\,(T\,(x))}{4v\,(T\,(x))}\right)^{\frac{1}{3}} n^{-\frac{1}{3}}, \tag{26}$$

where

$$u\,(T\,(x)) = m\,(y)\left(1 - \int_{-1}^{1} K^2\,(t)\,dt\right)$$

and

$$v\,(T\,(x)) = \left[\frac{1}{2} m'\,(y) \int_{-1}^{1} t^2 k\,(t)\,dt\right]^2.$$

For example, if we calculate $m\,(y)$ and $m'\,(y)$ at the 99% percentile of a $Beta\,(3,3)$ distribution, we find that

$$m\,(0.788\,72) = \frac{15}{16}\left(1 - 0.788\,72^2\right)^2 = 0.133\,90$$

and

$$m'\,(0.788\,72) = \left(\frac{15}{4} 0.788\,72\left(0.788\,72^2 - 1\right)\right)^2 = 1.\,249\,4.$$

Then:

$$b_{y=0.78872}^{Clas} = \left(\frac{0.133\,90\frac{9}{35}}{\left[\frac{1}{5}\right]^2 1.\,249\,4}\right)^{\frac{1}{3}} n^{-\frac{1}{3}} = 0.883\,21 n^{-\frac{1}{3}}$$

Following Bolancé et al. (2008a), we propose to implement double transformed kernel estimation of cdf as follows. First, use a transformation based on the generalized Champernowne distribution in (16) and then, use a Beta transformation. The quantile estimator is obtained from the estimated cdf.

Table 1: Distributions in the simulation study

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

6 Simulation Study

We summarize the results of a simulation study. We compare the MSE for estimating VaR_α of our proposed double transformed kernel estimation (DTKE), the empirical estimation (Emp), classical kernel estimation (CKE), kernel quantile estimation (KQE) and a transformed kernel estimation (TKE) using a transformation that is a cdf, namely the Champernowne cdf proposed in Buch-Larsen et al. (2005)[6].

We generated $2,000$ samples of size $n = 500$ and $2,000$ samples of size $n = 5,000$ from each distribution in Table 1. We selected four distributions with positive skewness and different tail shapes: Lognormal, Weibull and two mixtures of Lognormal-Pareto[7].

For each sample of size $n = 500$ we estimated the VaR_α, with $\alpha = 0.95$ and $\alpha = 0.995$. When the sample size is $n = 5,000$, in addition, we estimated VaR_α with $\alpha = 0.999$.

Using the $2,000$ replication estimates we estimated MSE for each method. To calculate MSE we used the theoretical value of VaR_α in Table 2. Results shown in Table 3 and Table 4 are the ratio between MSE of the different smoothing methods (KQE, CKE, TKE, DTKE) and the MSE of the empirical method (Emp). Moreover, sub-index w indicates that we used bandwidth based on asymptotic minimization of $WISE$ and sub-index x indicates that we used a smoothing parameter based on minimization of MSE for the corresponding α level.

In Table 3 we present the results for the Weibull and the Lognormal distributions. A value smaller than one indicates that the mean squared error is smaller than that of the empirical cdf

[6]We also studied the transformed kernel estimation of cdf proposed in Swanepoel and Van Graan (2005), but the results are worse than those obtained with the method of Buch-Larsen et al. (2005) and we do not include them in the tables with the rest of the simulation results. Moreover, we also simulated the TKE proposed in Bolancé et al. (2003), adapted to cdf estimation. There, a concave transformation is used that is not a cdf. We found that this method does not perform well for extreme quantiles and, therefore, we do not show these results in the summary tables.

[7]We used the same parameters as in Bolancé et al. (2008a).

Table 2: True VaR_α in the simulated distributions

Distribution	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Weibull	2.0781	3.0392	3.6271
Lognormal	2.2760	3.6252	4.6885
Mixture Lognormal-Pareto$_{p=0.7}$	7.5744	59.1892	299.0013
Mixture Lognormal-Pareto$_{p=0.3}$	13.4079	139.0034	699.0001

method. These results indicate that, when the true distribution does not have a heavy tail, i.e. Weibull or Lognormal, KQE and CKE provide good results and they are better than empirical cdf approach for all values of α. This improvement is larger when the sample size increases and α increases to 1.

In Table 4 we show the results of the ratio of MSE, for the two mixture Lognormal-Pareto distributions. These results show the advantages of kernel and transformed kernel estimation when the distribution has heavy tail. The results in Table 4 indicate that when the distribution has a very heavy tail the KQE and CKE do not outperform empirical cdf estimation approach (Emp), when $\alpha = 0.995$ and $\alpha = 0.999$, with both sample sizes $n = 500$ and $n = 5,000$.

The results in Table 4 show that the DTKE method improves the empirical cdf method, specially for a large sample size and an extreme quantile. For example, for a 70% Lognormal - 30% Pareto, and sample size $n = 5,000$, DTKE reduces the MSE of Emp in the estimation of the $VaR_{0.999}$ by 46%, i.e. the ratio between the MSE of DTKE and the MSE of Emp method is 0.54. We also see that DTKE reduces the MSE for the estimation of $VaR_{0.995}$ by 17% compared to the empirical method. For a 70% Lognormal - 30%, the reduction is 44% and 12%, respectively for $\alpha = 0.999$ and $\alpha = 0.995$. It is important to note that our proposal allows to calculate the asymptotically optimal bandwidth $b^{Clas}_{T(x)}$ in expression (26) without assuming a value for $T(x)$, given that we can calculate this exactly from the $Beta\,(3,3)$ distribution in expression (22). Finally, the results in the Appendix show that the DTKE method reduces the variance of Emp and, in some cases, increases the bias. However, the variance reduction compensates the increase of the bias and the MSE decreases.

Moreover, we see that the results of the transformed kernel estimation with boundary correction (TKE) cannot be used when estimating extreme quantiles. As the ratios are much above 1, we conclude that with this estimation method we could underestimate risk due to the large error. In Table 4, we can see that when sample size is $n = 500$, TKE improves the empirical cdf methods about 35% to 50% in terms of the amelioration of the MSE, but for $n = 5,000$ results for TKE are much worse. In the Appendix, we show details on the bias results for a

Table 3: Results for Weibull and Lognormal

	n=500		n=5000		
	Weibull				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	1	1	1	1	1
KQE	0.89	0.87	0.97	0.90	0.81
CKEw	0.89	0.94	0.97	0.93	0.85
CKEx	0.88	0.87	0.96	0.89	0.76
TKEw	0.75	10.31	0.90	19.60	28.08
DTKEw	0.95	2.15	0.98	1.29	1.42
DTKEx	0.92	1.25	0.97	0.98	0.89
	Lognormal				
	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	1	1	1	1	1
KQE	0.91	0.98	0.97	0.91	0.84
CKEw	0.92	0.98	0.97	0.94	0.93
CKEx	0.90	0.94	0.96	0.92	0.83
TKEw	0.65	8.90	0.87	24.76	24.31
DTKEw	0.94	2.07	0.96	1.22	1.25
DTKEx	0.92	1.24	0.95	0.97	0.83

Table 4: Results for mixture of Lognormal-Pareto

	n=500		n=5000		
	70% Lognormal-30% Pareto				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	1	1	1	1	1
KQE	1.03	294.29	0.97	1.06	647.32
CKEw	0.89	1.00	0.98	0.99	1.00
CKEx	0.90	1.00	0.98	0.99	1.00
TKEw	0.59	0.53	0.89	10.94	1.40
DTKEw	0.98	1.21	0.97	0.88	0.90
DTKEx	0.95	0.79	0.96	0.83	0.54
	30% Lognormal-70% Pareto				
	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	1	1	1	1	1
KQE	1.04	52.49	0.96	1.04	599.67
CKEw	0.91	1.00	0.98	1.00	1.00
CKEx	0.93	1.00	0.99	1.00	1.00
TKEw	0.49	0.65	0.86	12.04	1.58
DTKEw	0.89	1.90	0.93	0.98	1.07
DTKEx	0.87	1.12	0.93	0.88	0.66

closer inspection. There we can see that the TKE with boundary correction clearly underestimates the VaR_α. Therefore, we conclude that TKE is an unreliable and dangerous estimation procedure for extreme quantiles in risk management.

According to the simulation results and our theoretical approximations, we recommend to use a double transformation kernel estimation approximation with an optimal bandwidth to estimate VaR_α, for large databases and loss distributions that are heavy tailed.

An extension of this simulation study is reported in the Appendix. In order to analyze the sensibility of our method, we obtained additional simulation results by changing the theoretical parameters of the distribution in Table 1. The new distributions and their corresponding true VaR_α with the subsequent simulation results are shown in the Appendix. The new parameters cover a wide range of possible distributional tails. After carefully examining the results, we confirm the same conclusions mentioned before.

7 Conclusions

We have presented a method to estimate quantiles that is suitable when the loss is a random variable that is heavy tailed. The proposed double transformation kernel estimation does not depend on a parametric assumption for the random variable. Asymptotic properties have been proved, showing that when estimating extreme quantiles, the sample size needs to be large.

The proposed method is easily implemented and fast because the optimal smoothing parameter calculation is direct. Moreover, the proposed method is especially useful in many risk measurement settings because it does not require statistical distribution assumptions and can handle heavy tailed random variables. This is the case when analyzing some operational risk situations (see, Bolancé et al., 2012a) or in the analysis of severity distributions.

Our research provides a tractable nonparametric method that can be extended to other risk measures and can be useful to avoid restrictive statistical hypothesis.

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. We acknowledge Prof. Jens P. Nielsen for many fruitful discussions about the contents of this article. Montserrat Guillén thanks support from ICREA-Academia.

References

Altman, N., Léger, C., 1995. Bandwidth selection for kernel distribution function estimation. Journal of Statistical Planning and Inference 46, 195–214.

Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9, 203–228.

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., Alemany, R., Guillén, M., 2010a. Prediction of the economic cost of individual long-term care in the spanish population.

Bolancé, C., Guillén, M., Ayuso, M., 2012a. A nonparametric approach to analysing operational risk with an application to insurance fraud. The Journal of Operational Risk 7, 57–75.

Bolancé, C., Guillén, M., Gustafsson, J., Nielsen, J., 2012b. Quantitative Operational Risk Models. Chapman & Hall/CRC Finance Series, London.

Bolancé, C., Guillén, M., Nielsen, J., 2003. Kernel density estimation of actuarial loss functions. Insurance: Mathematics and Economics 32, 19–36.

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

Bolancé, C., Guillén, M., Nielsen, J., 2010b. Transformation kernel estimation of insurance claim cost distributions. Mathematical and Statistical Methods for Actuarial Sciences and Finance, 43–51.

Bolancé, C., Guillén, M., Pelican, E., Vernic, R., 2008b. Skewed bivariate models and nonparametric estimation for cte risk measure. Insurance: Mathematics and Economics 43, 386–393.

Bolancé, C., Guillén, M., Pinquet, J., 2008. On the link between credibility and frequency premium. Insurance: Mathematics and Economics 43 (2), 209–213.

Bowman, A., Hall, P., Prvan, T., 1998. Bandwidth selection for smoothing of distribution function. Biometrika 85, 799–808.

Buch-Kromann, T., Guillén, M., Linton, O., Nielsen, J., 2011. Multivariate density estimation using dimension reducing information and tail flattening transformations. Insurance: mathematics and economics 48 (1), 99–110.

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.

Cai, Z., Wang, X., 2008. Nonparametric estimation of conditional VaR and expected shortfall. Journal of Econometrics 147 (1), 120 – 130.

Dhaene, J., Vanduffel, S., Tang, Q., Goovaerts, M., Kaas, R., Vyncke, D., 2006. Risk measures and comonotonicity: A review. Stochastic Models 22, 573–606.

Eling, M., 2012. Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models? Insurance: Mathematics and Economics 51 (2), 239 – 248.

Englund, M., Guillen, M., Gustafsson, J., Nielsen, L., Nielsen, J., 2008. Multivariate latent risk: a credibility approach. Astin Bulletin 38 (1), 137–146.

Fan, J., Gu, J., 2003. Semiparametric estimation of value-at-risk. Econometrics Journal 6, 261–290.

Guillen, M., Gustafsson, J., Nielsen, J., Pritchard, P., 2007. Using external data in operational risk. The Geneva Papers on Risk and Insurance-Issues and Practice 32 (2), 178–189.

Guillén, M., Prieto, F., Sarabia, J., 2011. Modelling losses and locating the tail with the pareto positive stable distribution. Insurance: Mathematicsand Economics 49, 454–461.

Harrell, F., Davis, C., 1982. A new distribution-free quantile estimator. Biometrika 69, 635–640.

Hill, B., 1975. A simple general approach to inference about tail of a distribution. Annals of Statistics 3, 1163–1174.

Jones, B., Zitikis, R., 2007. Risk measures, distortion parameters, and their empirical estimation. Insurance: Mathematics and Economics 41 (2), 279 – 297.

Jorion, P., 2007. Value at Risk. McGraw-Hill, New York.

Kim, J., 2010. Bias correction for estimated distortion risk measure using the bootstrap. Insurance: Mathematics and Economics 47 (2), 198 – 205.

Krätschmer, V., Zähle, H., 2011. Sensitivity of risk measures with respect to the normal approximation of total claim distributions. Insurance: Mathematics and Economics 49 (3), 335 – 344.

Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives 3 (2), 73–84.

Lopez, O., 2012. A generalization of the kaplan-meier estimator for analyzing bivariate mortality under right-censoring and left-truncation with applications in model-checking for survival copula models. Insurance: Mathematics and Economics.

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

Peng, L., Qi, Y., Wang, R., Yang, J., 2012. Jackknife empirical likelihood method for some risk measures and related quantities. Insurance: Mathematics and Economics 51 (1), 142 – 150.

Pinquet, J., Guillén, M., Bolancé, C., 2001. Allowance for the age of claims in bonus-malus systems. Astin Bulletin 31 (2), 337–348.

Pitt, D., Guillén, M., Bolancé, C., 2012. An introduction to parametric and non-parametric models for bivariate positive insurance claim severity distributions. Xarxa de Referncia en Economia Aplicada (XREAP). Working Papers XREAP2010-03.

Reiss, R.-D., 1981. Nonparametric estimation of smooth distribution functions. Scandinavian Journal of Statistics 8, 116–119.

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

Ruppert, D. R., Cline, D. B. H., 1994. Bias reduction in kernel density estimation by smoothed empirical transformation. Annals of Statistics 22, 185–210.

Sarda, P., 1993. Smoothing parameter selection for smooth distribution functions. Journal of Statistical Planning and Inference 35, 65–75.

Sheather, S., Marron, J., 1990. Kernel quantile estimators. Journal of the American Statistical Association 85, 410–416.

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

Swanepoel, J., Van Graan, F., 2005. A new kernel distribution function estimator based on a nonparametric transformation of the data. Scandinavian Journal of Statistics 32, 551–562.

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.

Appendix

Proof 1 *Proof of Theorem 1:*

$$\begin{aligned}
E\left\{\widehat{F}_Y(y)-F_Y(y)\right\}^2 &\sim \frac{F_Y(y)\left[1-F_Y(y)\right]}{n}-f_Y(y)\frac{b}{n}\left(1-\int_{-1}^{1}K^2(t)\,dt\right)\\
&+\left[\frac{1}{2}f'_Y(y)\int_{-1}^{1}t^2k(t)\,dt\right]^2 b^4\\
&= \frac{F_Y(T(x))\left[1-F_Y(T(x))\right]}{n}-\frac{f_X(x)}{T'(x)}\frac{b}{n}\left(1-\int_{-1}^{1}K^2(t)\,dt\right)\\
&+\left[\frac{1}{2}\left(\frac{f_X(x)}{T'(x)}\right)'\int_{-1}^{1}t^2k(t)\,dt\right]^2 b^4\\
&= \frac{F_X(x)\left[1-F_X(x)\right]}{n}-\frac{1}{T'(x)}f_X(x)\frac{b}{n}\left(1-\int_{-1}^{1}K^2(t)\,dt\right)\\
&+\left[\frac{1}{2}\frac{f'_X(x)\,T'(x)-f_X(x)\,T''(x)}{\left(T'(x)\right)^2}\int_{-1}^{1}u^2k(t)\,dt\right]^2 b^4\\
&= \frac{F_X(x)\left[1-F_X(x)\right]}{n}-\frac{1}{T'(x)}f_X(x)\frac{b}{n}\left(1-\int_{-1}^{1}K^2(t)\,dt\right)\\
&+\left[\frac{1}{2}f'_X(x)\left[\frac{1}{T'(x)}-\frac{f_X(x)\,T''(x)}{f'_X(x)\left(T'(x)\right)^2}\right]\int_{-1}^{1}t^2k(t)\,dt\right]^2 b^4.
\end{aligned}$$

Table 5: Bias results mixture Lognormal-Pareto

	n=500		**n=5000**		
	70% Lognormal-30% Pareto				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	0.236	14.667	0.019	3.118	76.832
KQE	0.361	152.058	0.027	4.831	571.196
CKEw	0.130	14.671	0.009	2.880	76.568
CKEx	0.129	14.671	0.008	2.855	76.408
TKEw	0.031	-49.726	0.019	-43.516	-283.102
DTKEw	0.452	30.328	0.083	4.263	121.501
DTKEx	0.418	20.319	0.075	3.208	70.106
	30% Lognormal-70% Pareto				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	0.515	32.376	0.046	5.401	187.426
KQE	0.882	263.327	0.072	9.304	1143.475
CKEw	0.257	32.379	0.022	5.031	187.030
CKEx	0.278	32.378	0.021	5.047	186.883
TKEw	-0.365	-122.615	0.007	-107.893	-667.352
DTKEw	0.855	96.517	0.158	12.144	306.070
DTKEx	0.792	60.718	0.145	8.081	177.801

Table 6: Standard deviation results mixture Lognormal-Pareto

	n=500		**n=5000**		
	70% Lognormal-30% Pareto				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	1.147	66.591	0.337	12.785	226.987
KQE	1.135	1159.806	0.332	12.692	6070.150
CKEw	1.100	66.589	0.334	12.775	226.985
CKEx	1.106	66.588	0.335	12.773	226.983
TKEw	0.900	1.120	0.318	0.653	0.665
DTKEw	1.065	68.486	0.322	11.566	192.007
DTKEx	1.063	56.932	0.323	11.580	161.804
	30% Lognormal-70% Pareto				
Method	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Emp	2.889	148.576	0.834	30.624	495.938
KQE	2.854	1069.727	0.817	30.324	12932.464
CKEw	2.786	148.574	0.828	30.613	495.937
CKEx	2.815	148.574	0.829	30.615	495.937
TKEw	2.016	2.543	0.775	1.842	1.885
DTKEw	2.629	186.024	0.792	28.233	456.538
DTKEx	2.626	149.128	0.793	27.959	390.843

Table 7: New scenarios for parameter values in the extended simulation study for sensibility analysis

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

The remaining parameters are those already shown in Table 1.

Table 8: True VaR_α in the simulated distributions. Parameter scenario with values smaller than in Table 1)

Distribution	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Weibull	4.3185	9.2367	13.1558
Lognormal	1.5086	1.9040	2.1653
Mixture Lognormal-Pareto$_{p=0.7}$10	8.6258	93.6051	564.4016
Mixture Lognormal-Pareto$_{p=0.3}$10	18.0137	241.4306	1448.5061

Table 9: True VaR_α in the simulated distributions. Parameter scenario with values larger than in Table 1)

Distribution	$\alpha = 0.95$	$\alpha = 0.995$	$\alpha = 0.999$
Weibull	1.4416	1.7433	1.9045
Lognormal	5.1802	13.1422	21.9821
Mixture Lognormal-Pareto$_{p=0.7}$10	6.8606	40.9029	177.6320
Mixture Lognormal-Pareto$_{p=0.3}$10	10.5928	88.3539	384.8806

Table 10: Results for Weibull and Lognormal. Parameter scenario with values smaller than in Table 1)

	n=500		**n=5000**		
	Weibull				
Method	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	0.91	0.97	0.97	0.91	0.84
CKEw	0.92	0.99	0.98	0.95	0.93
CKEx	0.92	0.98	0.98	0.94	0.88
TKEw	0.78	7.40	0.90	10.45	17.17
DTKEw	1.02	2.92	1.00	1.48	1.62
DTKEx	0.98	1.55	0.99	1.05	0.95
	Lognormal				
	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	0.90	0.86	0.97	0.91	0.81
CKEw	0.91	0.97	0.97	0.94	0.91
CKEx	0.88	0.85	0.95	0.89	0.76
TKEw	0.68	10.83	0.87	27.39	31.13
DTKEw	0.96	1.91	0.96	1.20	1.23
DTKEx	0.94	1.17	0.95	0.96	0.82

Table 11: Results for Weibull and Lognormal. Parameter scenario with values larger than in Table 1)

	n=500		n=5000		
	Weibull				
Method	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	0.90	0.85	0.97	0.90	0.80
CKEw	0.91	0.91	0.97	0.92	0.83
CKEx	0.88	0.80	0.96	0.87	0.73
TKEw	0.79	8.08	0.91	1.92	16.36
DTKEw	0.98	2.21	0.99	1.37	1.40
DTKEx	0.95	1.26	0.98	1.00	0.88
	Lognormal				
	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	0.93	1.13	0.97	0.93	0.93
CKEw	0.90	0.99	0.97	0.95	0.97
CKEx	0.90	0.99	0.98	0.95	0.95
TKEw	0.65	5.66	0.87	20.23	15.01
DTKEw	1.00	2.41	0.97	1.27	1.30
DTKEx	0.97	1.37	0.96	0.99	0.84

Table 12: Results for mixture of Lognormal-Pareto. Parameter scenario with values smaller than in Table 1))

	n=500		**n=5000**		
	70% Lognormal-30% Pareto				
Method	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	1.09	4235.00	0.97	1.10	44063.73
CKEw	0.91	1.00	0.98	0.99	1.00
CKEx	0.92	1.00	0.98	1.00	1.00
TKEw	0.47	0.32	0.88	11.54	1.28
DTKEw	0.94	1.51	0.96	0.88	1.00
DTKEx	0.92	0.74	0.96	0.83	0.78
	30% Lognormal-70% Pareto				
	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	1.18	1143629.90	0.98	1.05	2460.04
CKEw	0.92	1.00	0.98	1.00	1.00
CKEx	0.94	1.00	0.98	1.00	1.00
TKEw	0.51	0.23	0.85	9.76	1.04
DTKEw	0.90	1.02	0.94	0.97	1.36
DTKEx	0.89	0.57	0.94	0.85	0.72

Table 13: Results for mixture of Lognormal-Pareto. Parameter scenario with values larger than in Table 1)

	n=500		**n=5000**		
	70% Lognormal-30% Pareto				
Method	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	1.00	272.13	0.97	1.04	911.01
CKEw	0.91	1.00	0.98	0.99	1.00
CKEx	0.91	1.00	0.98	0.99	1.00
TKEw	0.68	0.60	0.88	11.71	2.17
DTKEw	0.99	0.95	0.96	0.88	1.00
DTKEx	0.97	0.70	0.95	0.84	0.63
	30% Lognormal-70% Pareto				
	a=0.95	a=0.995	a=0.95	a=0.995	a=0.999
Emp	1	1	1	1	1
KQE	1.04	24357.97	0.97	1.00	95.70
CKEw	0.92	1.00	0.98	0.99	1.00
CKEx	0.93	1.00	0.98	0.99	1.00
TKEw	0.61	0.57	0.87	11.47	1.88
DTKEw	0.93	1.06	0.95	0.93	0.99
DTKEx	0.91	0.70	0.95	0.85	0.67

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)

XARXA DE REFERÈNCIA EN ECONOMIA APLICADA

xreap@pcb.ub.es