


RESEARCH ARTICLE

Estimating the VaR-induced Euler allocation rule

N.V. Gribkova^{1,2}, J. Su^{3,4}  and R. Zitikis^{4,5}

¹Saint Petersburg State University, Saint Petersburg, 199034 Russia, ²Emperor Alexander I St.Petersburg State Transport University, Saint Petersburg, 190031 Russia, ³Purdue University, West Lafayette, IN 47907, USA, ⁴Risk and Insurance Studies Centre, York University, Toronto, Ontario M3J 1P3, Canada and ⁵Western University, London, Ontario N6A 5B7, Canada
Corresponding author: J. Su; Email: jianxi@purdue.edu

Received: 22 September 2022; **Revised:** 12 March 2023; **Accepted:** 29 March 2023; **First published online:** 2 May 2023

Keywords: Capital allocations; conditional mean risk sharing; quantile regression; order statistics; concomitants.

Abstract

The prominence of the Euler allocation rule (EAR) is rooted in the fact that it is the only return on risk-adjusted capital (RORAC) compatible capital allocation rule. When the total regulatory capital is set using the value-at-risk (VaR), the EAR becomes – using a statistical term – the quantile-regression (QR) function. Although the cumulative QR function (i.e., an integral of the QR function) has received considerable attention in the literature, a fully developed statistical inference theory for the QR function itself has been elusive. In the present paper, we develop such a theory based on an empirical QR estimator, for which we establish consistency, asymptotic normality, and standard error estimation. This makes the herein developed results readily applicable in practice, thus facilitating decision making within the RORAC paradigm, conditional mean risk sharing, and current regulatory frameworks.

1. Introduction

The Euler allocation rule (EAR) plays a unique role in financial and insurance risk management as it is the only return on risk-adjusted capital (RORAC) compatible capital allocation rule (e.g., Tasche, 2007; McNeil *et al.*, 2015, and references therein). In the currently adopted regulatory frameworks, such as Basel II and III, Solvency II, and the Swiss Solvency Test (EIOPA, 2016; IAIS, 2016; BCBS, 2016, 2019), the value-at-risk (VaR) and the expected shortfall (ES) play prominent roles.

The ES-induced EAR is the tail conditional allocation (TCA), whose empirical estimation from several perspectives and under minimal conditions has recently been developed by Gribkova *et al.* (2022a, b), using the therein proposed technique that hinges on compound sums of concomitants. They assume that data arise from independent and identically distributed (iid) random variables. For a much wider class of data generating process such as time series, albeit under conditions that are stronger in the iid case than those imposed in the aforementioned two papers, we refer to Asimit *et al.* (2019), where earlier references on the topic can also be found. We shall employ certain elements of the technique of Gribkova *et al.* (2022a,b), in the current paper as well, crucially supplementing it with kernel-type smoothing. Note also that the TCA is a special case of the weighted capital allocation (Furman and Zitikis 2008, 2009) for which an empirical estimation theory has been developed by Gribkova and Zitikis (2017, 2019), where extensive references to the earlier literature can also be found.

We are grateful to two anonymous reviewers and the editor in charge of the manuscript for their careful analyses of our theoretical and numerical results, constructive criticism, suggestions, and guidance. This research has been supported by the NSERC Alliance-Mitacs Accelerate grant entitled “New Order of Risk Management: Theory and Applications in the Era of Systemic Risk” from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the national research organization Mathematics of Information Technology and Complex Systems (MITACS) of Canada.

In the present paper, we lay statistical foundations for estimating the VaR-induced EAR, which is known as the conditional mean risk sharing in peer-to-peer (P2P) insurance, a topic of much recent interest among insurance practitioners and academics. For a few references on the topic, we refer to Denuit (2019) and Denuit and Robert (2020), whereas Abdikerimova and Feng (2022) and Denuit *et al.* (2022) provide a wealth of information about the P2P insurance. In particular,

- using the notation of concomitants, we define a consistent empirical EAR estimator,
- establish its asymptotic normality under minimal conditions,
- propose a consistent empirical estimator of the standard VaR-induced EAR error.

These contributions make the herein developed theory readily applicable in practice, including constructing confidence intervals for, and testing hypotheses about, the VaR-induced EAR. We shall illustrate numerical performance of these results later in this paper.

In detail, for a given probability level $p \in (0, 1)$ and a risk variable Y , whose cumulative distribution function (cdf) we denote by G , the VaR is given by

$$\text{VaR}_p(Y) = \inf\{y \in \mathbb{R} : G(y) \geq p\}. \quad (1.1)$$

In the statistical parlance, $\text{VaR}_p(Y)$ is the p th quantile of Y , usually denoted by $G^{-1}(p)$, and in the mathematical parlance, it is the left-continuous inverse of G .

Financial and insurance institutions usually have several business lines. When the total capital is calculated using VaR, the allocated capital to the business line with risk X is, according to the EAR (e.g., Tasche, 2007; McNeil *et al.*, 2015), given by

$$\text{EAR}_p(X | Y) = \mathbb{E}(X | Y = \text{VaR}_p(Y)).$$

A clarifying note is now warranted.

Note 1.1. As an offspring of econometric and statistical problems, $\text{EAR}_p(X | Y)$ has appeared in the statistical literature under the name of quantile regression function (e.g., Rao and Zhao, 1995; Tse, 2009, and references therein), which stems from the fact that it is the composition $r_{X|Y} \circ G^{-1}(p)$ of the least-squares regression function $r_{X|Y}$ and the quantile function G^{-1} . Note in this regard that $\text{EAR}_p(X | Y)$, despite being called the quantile regression function by statisticians, is only superficially connected to the research area commonly known as Quantile Regression (Koenker 2005).

Naturally, it is desirable to empirically estimate $\text{EAR}_p(X | Y)$. For the empirical EAR estimator that we shall formally introduce in the next section, we have developed a thorough statistical inference theory under minimal conditions. Namely, in Section 2 we define the estimator and also its asymptotically equivalent version to facilitate diverse computing preferences. In the same section, we present three theorems that are the main results of this paper: consistency of the EAR estimator, its asymptotic normality, and standard error. These results are illustrated using simulated data in Section 3, and then applied on real data in Section 4. Section 5 concludes the paper. Proofs and technical lemmas are in the Online Supplement (see Section S1).

2. Estimators and their large-sample properties

We start with several basic conditions on the distribution of the pair (X, Y) , whose joint cdf we denote by H . The marginal cdf's of X and Y are denoted by F and G , respectively. We have already introduced the left-continuous inverse $p \mapsto \text{VaR}_p(Y)$ of the cdf G . The right-continuous inverse $p \mapsto \text{V@R}_p(Y)$ is defined by

$$\text{V@R}_p(Y) = \sup\{y \in \mathbb{R} : G(y) \leq p\}.$$

Next are the first two conditions that we impose throughout the paper.

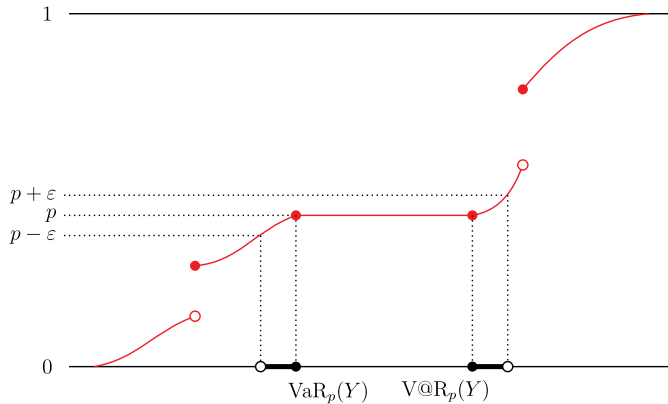


Figure 1. The cdf G is continuous in a neighborhood of the interval $[VaR_p(Y), V@R_p(Y)]$ and strictly increasing a bit to the left of $VaR_p(Y)$ and a bit to the right of $V@R_p(Y)$.

(C₁) There exists $\varepsilon > 0$ such that the cdf G is continuous and strictly increasing on the set

$$V_\varepsilon = (VaR_{p-\varepsilon}(Y), VaR_p(Y)] \cup [V@R_p(Y), V@R_{p+\varepsilon}(Y)).$$

(C₂) The function $\tau \mapsto EAR_\tau(X | Y) = g(VaR_\tau(Y))$ is finite and continuous in a neighborhood of p , where g is the regression function

$$g(y) = \mathbb{E}(X | Y = y).$$

Figure 1 illustrates condition (C₁). Note that the gap $(VaR_p(Y), V@R_p(Y))$ between the two intervals in the definition of V_ε coincides with the interval where the random variable Y does not (almost surely) place any values, and thus conditioning on such values does not make sense. Indeed, all conditional expectations are defined only almost surely.

Note 2.1. The reason we need the continuities specified in conditions (C₁) and (C₂) is that $EAR_p(X | Y)$ is the regression function $g(y)$ evaluated at the single point $y = VaR_p(Y)$, and thus to gather sufficient information (that is, data) about the EAR we need to combine the information from neither too large nor too small neighborhoods of the point y , the latter being of course unknown. From this perspective, conditions (C₁)–(C₂) are natural and basic.

Let $(X_1, Y_1), (X_2, Y_2), \dots$ be a sequence of independent copies of the pair (X, Y) , and let G_n denote the empirical cdf based on Y_1, \dots, Y_n , that is,

$$G_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i \leq y\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_{i:n} \leq y\}},$$

where $\mathbb{1}$ is the indicator function, and $Y_{1:n} \leq \dots \leq Y_{n:n}$ are the order statistics of Y_1, \dots, Y_n . (In the case of possible ties, the order statistics of Y_1, \dots, Y_n are enumerated arbitrarily.) An empirical estimator for $VaR_p(Y)$ can be constructed by replacing G by G_n in Equation (1.1). In a computationally convenient form, the obtained estimator can be written as

$$VaR_{p,n} = Y_{\lceil np \rceil, n},$$

where $\lceil \cdot \rceil$ is the ceiling function.

Since the order statistics that we need for estimating $EAR_p(X | Y)$ can lie on either side of $VaR_{p,n}$, we control their locations using two sequences of non-negative real numbers: $\Delta_{1,n}$ and $\Delta_{2,n}$ for $n \geq 1$. They define neighborhoods of $VaR_{p,n}$ from which we collect the order statistics needed for the construction of an empirical EAR estimator. The two sequences need to satisfy the following conditions:

(D₁) $\max(\Delta_{1,n}, \Delta_{2,n}) \rightarrow 0$ as $n \rightarrow \infty$.

(D₂) $\liminf_{n \rightarrow \infty} \sqrt{n}(\Delta_{1,n} + \Delta_{2,n}) > 0$.

Note 2.2. The appearance of these Δ 's (a.k.a. band widths) when estimating quantities that involve conditioning on events of zero probabilities (as is the case with the VaR-induced EAR) is unavoidable, theoretically speaking. In practice, however, when working with real data sets, which are concrete and of fixed sizes, the choices of Δ 's can be data driven, as it would be when using, for example, cross-validation (e.g., James *et al.*, 2013). We do not follow this path in the current paper, given that we have developed satisfactory intuitive understanding of what appropriate Δ 's could be, with details in Section 3.2 below.

Hence, according to conditions (D₁)–(D₂), the aforementioned neighborhoods should shrink, but not too fast. The estimator of $\text{EAR}_p(X|Y)$ that we shall formally define in a moment is based on averaging those X 's whose corresponding Y 's are near, as determined by the two Δ 's, the order statistic $Y_{[np]:n}$. To formalize the idea, we first recall that the induced by Y_1, \dots, Y_n order statistics, usually called concomitants, are the first coordinates $X_{1,n}, \dots, X_{n,n}$ when, without changing the composition of the pairs $(X_1, Y_1), \dots, (X_n, Y_n)$, we order them in such a way that the second coordinates become ascending, thus arriving at the pairs $(X_{1,n}, Y_{1,n}), \dots, (X_{n,n}, Y_{n,n})$. The empirical EAR estimator is now given by the formula

$$\text{EAR}_{p,n} = \frac{1}{N_n} \sum_{i=1}^n X_{i,n} \mathbb{1}_{(p-\Delta_{1,n}, p+\Delta_{2,n})} \left(\frac{i}{n} \right), \quad (2.1)$$

where the indicator $\mathbb{1}_{(a,b)}(t)$ is equal 1 when $t \in (a, b)$ and 0 otherwise, and

$$N_n = \sum_{i=1}^n \mathbb{1}_{(p-\Delta_{1,n}, p+\Delta_{2,n})} \left(\frac{i}{n} \right).$$

A few clarifying notes are in order.

Note 2.3. In the case of possible ties, the order statistics of Y_1, \dots, Y_n are enumerated arbitrarily. This arbitrariness does not affect the EAR estimator due to condition (C₁). Indeed, $\text{EAR}_{p,n}$ is based on the Y -order statistics in shrinking neighborhoods of the point $\text{VaR}_p(Y)$, around which the cdf G is continuous by condition (C₁), and thus the Y 's that fall into the neighborhoods are (almost surely) distinct, thus giving rise to the uniquely defined concomitants that are not nullified by the indicators on the right-hand side of Equation (2.1).

Note 2.4. It would not be right for us to claim that the estimator $\text{EAR}_{p,n}$ is new, as in various guises it has appeared in research by, for example, Gouriéroux *et al.* (2000), Tasche (2007, 2009), Fu *et al.* (2009), Hong (2009), Liu and Hong (2009), and Jiang and Fu (2015). What is new in the current formulation of the estimator is the introduction of the notion of concomitants, which open up the doors into a vast area of Statistics associated with order statistics, concomitants, ranks, and other related objects whose properties have been extensively investigated (e.g., David and Nagaraja, 2003). It is this knowledge coupled with methodological inventions of Borovkov (1988) that has allowed us to establish the results of the present paper under minimal conditions.

Note 2.5. Deriving statistical inference for the VaR-induced EAR is much more difficult than doing the same for the ES-induced EAR, which explains why there is always a time lag before we see results for the VaR-induced EAR in the literature. Very interestingly, we note in this regard that Asimit *et al.* (2019) observed that under some conditions and for large values of p (that is, for those close to 1), the VaR-induced EAR and the ES-induced EAR are close to each other, thus helping to circumvent the challenges associated with the VaR-induced EAR, given that statistical inference for the ES-induced EAR is easier, and is often already available in the literature. This trick, however, comes at an expense associated with a specific choice of p , which depends on the population distribution and thus needs to be estimated. For details, we refer to Asimit *et al.* (2019).

We are now ready to formulate our first theorem concerning $\text{EAR}_{p,n}$.

Theorem 2.1. *When conditions (C₁)–(C₂) and (D₁)–(D₂) are satisfied, the empirical EAR estimator consistently estimates EAR_p(X | Y), that is,*

$$\text{EAR}_{p,n} \xrightarrow{P} \text{EAR}_p(X | Y) \tag{2.2}$$

when $n \rightarrow \infty$.

Given the definition of EAR_p(X | Y), the estimator EAR_{p,n} is highly intuitive and is therefore used in all the theorems of this paper, but from the computational point of view, we may find other asymptotically equivalent versions more convenient. One of them, which we shall also use in our numerical studies as well as in some proofs, is given by

$$\widehat{\text{EAR}}_{p,n} = \frac{1}{k_{2,n} - k_{1,n} + 1} \sum_{i=k_{1,n}}^{k_{2,n}} X_{i,n}, \tag{2.3}$$

where the integers $k_{1,n}$ and $k_{2,n}$ are defined by

$$k_{1,n} = [n(p - \Delta_{1,n})] \quad \text{and} \quad k_{2,n} = [n(p + \Delta_{2,n})] \tag{2.4}$$

with $[\cdot]$ denoting the greatest integer function. Theorem 2.1 as well as the two theorems that we shall formulate later in this section hold if we replace EAR_{p,n} by $\widehat{\text{EAR}}_{p,n}$.

Note 2.6. We have achieved the minimality of conditions in all our results by crucially exploiting distributional properties of concomitants. Therefore, the estimator $\widehat{\text{EAR}}_{p,n}$ frequently turns out to be more convenient than EAR_{p,n}. Nevertheless, asymptotically when the sample size n increases, the two empirical estimators are equivalent. There are, however, lines of research when EAR_{p,n} becomes more convenient. For example, it allows to naturally introduce kernel-type smoothing, which in the technical language means replacing the indicator on the right-hand side of Equation (2.1) by a kernel function (e.g., Silverman, 1986). Although this is a very useful and interesting research, we have refrained from it in the present paper as, inevitably, conditions on the kernel function need to be imposed, and they will interact with other conditions such as those on the bandwidth and the population distribution, thus impeding our plan to develop statistical inference for EAR under minimal conditions.

Conditions (C₁)–(C₂) and (D₁)–(D₂) imposed for the first-order result (i.e., consistency) also give clues as to what would be needed for second-order results, such as asymptotic normality and standard error estimation.

- (C₃) The function $\tau \mapsto \text{EAR}_\tau(X | Y)$ is finite and α -Hölder continuous for some $\alpha \in (1/2, 1]$ at the point p , that is, there exists a neighborhood of the point p and also a constant $L \in (0, \infty)$ such that the bound

$$|\text{EAR}_\tau(X | Y) - \text{EAR}_p(X | Y)| \leq L|\tau - p|^\alpha$$

holds for all τ in the aforementioned neighborhood.

- (C₄) The function $\tau \mapsto \mathbb{E}(X^2 | Y = \text{VaR}_\tau(Y)) = g_2(\text{VaR}_\tau(Y))$ is finite and continuous in a neighborhood of p , where the function g_2 is defined by

$$g_2(y) = \mathbb{E}(X^2 | Y = y).$$

Note 2.7. The verification of condition (C₃) is not expected to pose serious obstacles in practice as it is satisfied when the first derivative of the function $\tau \mapsto \text{EAR}_\tau(X | Y)$ is bounded in a neighborhood of p . This means that the function should not have a jump at the point p . Naturally, practitioners have good intuition whether this assumption is plausible, given their subject matter knowledge and, in particular, their chosen model for (X, Y) .

While conditions (D₁)–(D₂) require the two Δ 's to converge to 0 due to EAR being the regression function evaluated at a single point, the convergence should not be too fast in order to enable the collection of sufficiently many data points. On the other hand, intuitively, a valid asymptotic normality result

should require the two Δ 's to converge to 0 fast enough to avoid creating a bias. This is the reason behind the following condition:

$$(D_3) \quad n^{1/(2\alpha+1)} \max(\Delta_{1,n}, \Delta_{2,n}) \rightarrow 0 \text{ when } n \rightarrow \infty.$$

Note 2.8. It is condition (D_3) that has forced the restriction $\alpha > 1/2$ in condition (C_3) . In the special case $\alpha = 1$, which corresponds to the case when $\tau \mapsto \text{EAR}_\tau(X | Y)$ is a Lipschitz continuous function, condition (D_3) reduces to $n^{1/3} \max(\Delta_{1,n}, \Delta_{2,n}) \rightarrow 0$; compare it with condition (D_2) .

We are now ready to formulate our asymptotic normality result for the empirical EAR estimator, with \mathbb{V} denoting the variance operator.

Theorem 2.2. *When conditions (C_1) – (C_4) and (D_1) – (D_3) are satisfied, the empirical EAR estimator is asymptotically normal, that is,*

$$\sqrt{N_n}(\text{EAR}_{p,n} - \text{EAR}_p(X | Y)) \xrightarrow{d} \mathcal{N}(0, \sigma^2), \tag{2.5}$$

where σ^2 is the asymptotic variance given by the formula

$$\sigma^2 = \mathbb{V}(X | Y = \text{VaR}_p(Y)). \tag{2.6}$$

To give an insight into the normalizing factor $\sqrt{N_n}$, which should not, of course, be confused with \sqrt{n} , we let

$$\Delta_{1,n} = \Delta_{2,n} =: \Delta_n$$

and then set

$$\Delta_n = n^{-\eta}$$

with various parameter $\eta > 0$ values to be discussed next. Conditions (D_1) – (D_2) are satisfied if and only if $\eta \in (0, 1/2]$. In this case, we have $N_n \sim n^{1-\eta}$, with the familiar in density estimation cases $N_n \sim n^{4/5}$ and $N_n \sim n^{2/3}$ when $\eta = 1/5$ and $\eta = 1/3$, respectively. It is useful to note the following relationships between these values of η and the parameter α in condition (D_3) :

- if $\eta = 1/2$, then $\alpha > 1/2$;
- if $\eta = 1/3$, then $\alpha > 1$, which is not possible due to condition (C_3) ;
- if $\eta = 1/5$, then $\alpha > 2$, which is not possible due to condition (C_3) .

According to condition (C_3) , we must have $\alpha \in (1/2, 1]$ and thus the admissible range of η values becomes the interval

$$\left(\frac{1}{1+2\alpha}, \frac{1}{2} \right] \subseteq \left(\frac{1}{3}, \frac{1}{2} \right]$$

which reduces to the maximal interval $(1/3, 1/2]$ when $\alpha = 1$, that is, when the function $\tau \mapsto \text{EAR}_\tau(X | Y)$ is 1-Hölder (i.e., Lipschitz) continuous. We shall need to keep this in mind when setting parameters in the following simulation study, and also when studying a real data set later in this paper.

Hence, in view of the above, we conclude that, within the framework of the previous paragraph, the normalizer $\sqrt{N_n}$ in Theorem 2.2 is, asymptotically,

$$\sqrt{N_n} \sim n^{\eta^*}$$

with the parameter

$$\eta^* := \frac{1-\eta}{2} \in \left[\frac{1}{4}, \frac{1}{2+\alpha^{-1}} \right) \subseteq \left[\frac{1}{4}, \frac{1}{3} \right)$$

that can get arbitrarily close to $1/3$ from below if $\tau \mapsto \text{EAR}_\tau(X | Y)$ is 1-Hölder (i.e., Lipschitz) continuous, and it can be as low as $\eta^* = 1/4$ when $\eta = 1/2$. The radius (i.e., a half of the width) of the neighborhood of p from which the data are collected is then equal to

$$\Delta_n = n^{2\eta^*-1},$$

and thus the smoother the function $\tau \mapsto \text{EAR}_\tau(X | Y)$ is (i.e., smaller $\alpha > 0$), the narrower the neighborhood can be chosen. Hence, the normalizer $\sqrt{N_n}$ in Theorem 2.2 cannot become \sqrt{n} , a fact already noted by Gouriéroux *et al.* (2000), Hong (2009), and Liu and Hong (2009).

For practical purposes, we need an estimator of the variance σ^2 . We define it in the following theorem, where we also show that the estimator is consistent. (We use $:=$ when wishing to emphasize that certain equations hold by definition, rather than by derivation.)

Theorem 2.3. *When conditions (C₁)–(C₄) and (D₁)–(D₃) are satisfied, we have*

$$\widehat{\sigma}_{p,n}^2 := \frac{1}{N_n} \sum_{i=1}^n X_{i,n}^2 \mathbb{1}_{(p-\Delta_{1,n}, p+\Delta_{2,n})} \left(\frac{i}{n} \right) - (\text{EAR}_{p,n})^2 \xrightarrow{\mathbb{P}} \sigma^2. \tag{2.7}$$

With the help of classical Slutsky’s arguments, Theorems 2.2 and 2.3 immediately imply

$$\sqrt{\frac{N_n}{\widehat{\sigma}_{p,n}^2}} (\text{EAR}_{p,n} - \text{EAR}_p(X | Y)) \xrightarrow{d} \mathcal{N}(0, 1), \tag{2.8}$$

which enables to construct confidence intervals for, and test hypotheses about, $\text{EAR}_p(X | Y)$.

We conclude this section by reflecting upon the main results and their possible extensions, or generalizations. To begin with, recall that we have introduced and used the estimators $\text{EAR}_{p,n}$ and $\widehat{\text{EAR}}_{p,n}$, which rely on the bandwidths $\Delta_{1,n}$ and $\Delta_{2,n}$. They are the only parameters that require tuning. The simplicity of the VaR-induced EAR estimators has allowed us to derive their consistency and asymptotic normality under particularly weak conditions, virtually only under those that are required for the existence of the quantities involved in the formulations of Theorems 2.1–2.3. Yet, the conditions are plentiful, seven in total. Nevertheless, we can already start contemplating of extending these results in several directions.

For example, we can replace the indicators in the EAR and σ^2 estimators by some kernel functions, as it is done in the classical kernel-type density estimation (e.g., Silverman, 1986). This would of course necessitate the introduction of a set of conditions on the kernel functions, which would naturally be tied to the choices of $\Delta_{1,n}$ and $\Delta_{2,n}$.

One may also explore other EAR estimators, such as (recall Note 1.1) the one that arises by replacing the least-squares regression function $r_{X|Y}$ and the quantile function G^{-1} by their empirical estimators, such as the Nadaraya-Watson (or some other) estimator for the regression function $y \mapsto r_{X|Y}(y)$ and the empirical (or smoothed) estimator for the quantile function $G^{-1}(p)$. Naturally, we would inevitably be facing the necessity of a set of assumptions on the kernel functions used in the construction of the Nadaraya-Watson or other kernel-type estimators. We wish to note at this point, however, that the combined set of conditions arising from such estimators and their asymptotic theories are more stringent than the conditions of Theorems 2.1–2.3. This is natural because our proofs have been designed to tackle the empirical EAR-estimators $\text{EAR}_{p,n}$ and $\widehat{\text{EAR}}_{p,n}$ as the whole and not componentwise.

3. The estimator in simulated scenarios

We shall now illustrate the performance of the empirical EAR estimator with the aid of simulated and real data. The section consists of two parts. In Section 3.1, we introduce an insurance-inspired model for simulations and use thus obtained data to evaluate the estimator. In Section 3.2, we work out intuition on bandwidth selection for practical purposes, which we later employ for analyzing a real data set in Section 4.

3.1. Simulation setup and the estimator’s performance

To facilitate a comparison between our earlier and current inference results, we adopt the same setup for simulations as in Gribkova *et al.* (2022b). Namely, we hypothesize a multiple-peril insurance product that contains two coverages with their associated losses L_1 and L_2 that follow the bivariate Pareto

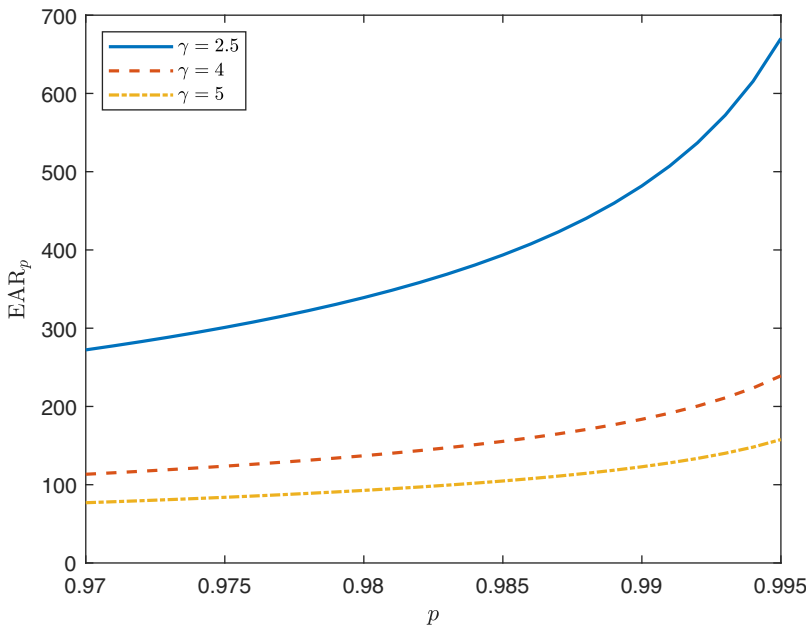


Figure 2. The plot of $EAR_{\tau}(X | Y)$ as a function of $\tau \in (0.975, 0.990)$ for distribution (3.1) with varying values of the tail parameter γ .

distribution whose joint survival function is

$$S(l_1, l_2) = (1 + l_1/\theta_1 + l_2/\theta_2)^{-\gamma}, \quad l_1 > 0, l_2 > 0, \tag{3.1}$$

where $\theta_1 > 0$ and $\theta_2 > 0$ are scale parameters, and $\gamma > 0$ is a shape parameter. (Note that smaller values of γ correspond to heavier tails.) Moreover, we assume that the two coverages have deductibles $d_1 > 0$ and $d_2 > 0$, and so the insurance payments are

$$W_i = (L_i - d_i) \times \mathbb{1}_{\{L_i > d_i\}}.$$

We are interested in evaluating the risk contribution of the first coverage out of the total loss using $EAR_p(X | Y)$ with $X = W_1$ and $Y = W_1 + W_2$. For various applications of the distribution, we refer to Alai et al. (2016), Su and Furman (2016), Sarabia et al. (2018), and references therein.

We set the parameters of the bivariate Pareto distribution (3.1) to $\theta_1 = 100$, $\theta_2 = 50$, and $\gamma \in \{2.5, 4, 5\}$, which are along the lines of the Pareto distribution’s parameter choices when modeling dependent portfolio risks in Su and Furman (2016). The deductibles are assumed to be $d_1 = 18$ and $d_2 = 9$, which are near the medians of L_1 and L_2 , respectively, when $\gamma = 4$. Using a formula provided by Gribkova et al. (2022b) for computing $EAR_p(X | Y)$ in the case of distribution (3.1), we compute $EAR_p(X | Y)$ for the parameter choices $p = 0.975$ and $p = 0.990$, which are motivated by the confidence levels considered in the currently adopted regulatory frameworks.

We depict $EAR_{\tau}(X | Y)$ as a function of τ with varying γ ’s in Figure 2. We see from it that the derivative of $\tau \mapsto EAR_{\tau}$ is finite at any point $\tau \in (0.975, 0.990)$. Thereby, condition (C₃) holds at least when $\alpha = 1$. For this reason, we set the bandwidths to

$$\Delta_{1,n} = \Delta_{2,n} = n^{-1/2}, \tag{3.2}$$

which ensures that all the conditions of our main results are satisfied.

Next, we pretend that the population distribution is unknown, as it would be in practice, and employ $\widehat{EAR}_{p,n}$ to estimate $EAR_p(X | Y)$. To demonstrate the large-sample precision of the estimator, we first consider a situation in which the user can generate data sets of any size (think, for example, of Economic

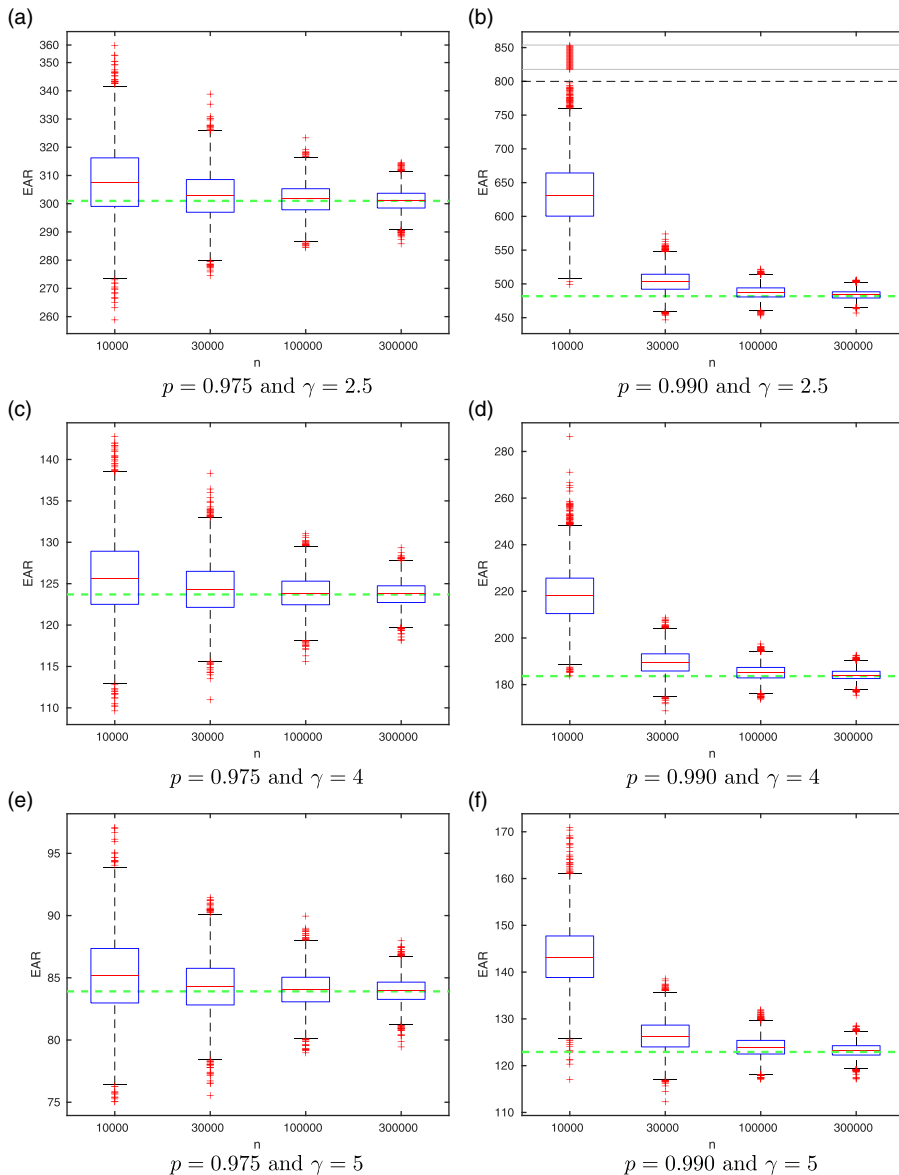


Figure 3. Box plots of the EAR estimates with varying values of the parameter γ and the sample size n , with the true $EAR_p(X|Y)$ depicted as the horizontal dashed green line.

Scenario Generators). For fixed sample sizes $n = i \times 10,000, i \in \{1, 3, 10, 30\}$, the same simulation exercise is repeated 50,000 times in order to assess the variability of $\widehat{EAR}_{p,n}$. The EAR estimates are summarized using box plots in Figure 3.

Note 3.1. The two vertical lines at the top of panel (b) in Figure 3 are due to extreme-data compression: “If any data values fall outside the limits specified by ‘DataLim’, then boxplot displays these values evenly distributed in a region just outside DataLim, retaining the relative order of the points.” (ExtremeMode, 2023)

We see from Figure 3 that among all the cases, the estimates produced by $\widehat{EAR}_{p,n}$ converge to the true value of $EAR_p(X|Y)$ when the sample size n grows, with the distribution of $\widehat{EAR}_{p,n}$ becoming narrower.

Table 1. Coverage percentages of the 90%-level confidence interval (3.3) for $EAR_p(X | Y)$ based on 50,000 sets of simulated data for each sample size n .

n	$p = 0.975$			$p = 0.990$		
	$\gamma = 2.5$	$\gamma = 4$	$\gamma = 5$	$\gamma = 2.5$	$\gamma = 4$	$\gamma = 5$
1×10^4	68%	72%	73%	1%	2%	3%
3×10^4	76%	79%	79%	41%	52%	55%
1×10^5	82%	84%	84%	69%	75%	75%
3×10^5	85%	86%	86%	79%	82%	83%

When $p = 0.975$, the intervals of the box plots cover the true value of EAR for all the selected n 's, although this is not the case when $p = 0.990$. This indicates that larger p 's diminish the performance of $\widehat{EAR}_{p,n}$, which is natural. On the other hand, the parameter γ , which governs the tail behavior of distribution (3.1), does not significantly impact the performance of $\widehat{EAR}_{p,n}$, although if we compare the cases $\gamma = 2.5$ and $\gamma = 5$, the estimator seems to behave slightly better when γ is larger.

In addition to the above discussion based on box plots, we shall now use the same simulation design to calculate coverage proportions of the parameter $EAR_p(X | Y)$ by the $100 \times (1 - \nu)\%$ -level confidence interval

$$\left(\widehat{EAR}_{p,n} - z_{\nu/2} \widehat{\sigma}_{n,p} / \sqrt{N_n}, \widehat{EAR}_{p,n} + z_{\nu/2} \widehat{\sigma}_{n,p} / \sqrt{N_n} \right), \tag{3.3}$$

where $z_{\nu/2}$ is the z -value with $\nu \in (0, 1)$ set to 0.1. Hence, specifically, we count the coverage proportions of the 90%-level confidence intervals based on 50,000 sets of simulated data, repeating the same procedure for the sample sizes $n \in \{1, 3, 10, 30\} \times 10,000$. The results are summarized in Table 1. We see from the table that when n is small, p is large, and the tail of distribution (3.1) is heavy, confidence interval (3.3) barely captures the true EAR value. We have already encountered this feature in Figure 3, which reflects the fact that the bias of $\widehat{EAR}_{p,n}$ can be large under the aforementioned challenging data-generating scenario. However, as the sample size n increases, the bias diminishes considerably, and so the coverage proportions tend to 0.9 across the considered scenarios.

The distribution of $\widehat{EAR}_{p,n}$ or, more precisely, the distribution of its centered counterpart

$$\widehat{EAR}_{p,n}^* := \widehat{EAR}_{p,n} - \text{Average}(\widehat{EAR}_{p,n})$$

plays an important role when constructing confidence intervals for, and hypothesis tests about, $EAR_p(X | Y)$. When we can simulate data sets of arbitrary size, the distribution of $\widehat{EAR}_{p,n}^*$ can be obtained via Monte Carlo, but in most practical cases, we only have one data set. In such cases, Theorem 2.2 says that we can approximate the distribution of $\widehat{EAR}_{p,n}^*$ by $\mathcal{N}(0, \widehat{\sigma}_{p,n}^2 / N_n)$, where $\widehat{\sigma}_{p,n}^2$ comes from Theorem 2.3. The plots in Figure 4 compare the cdf of $\widehat{EAR}_{p,n}^*$ obtained via Monte Carlo with the cdf of $\mathcal{N}(0, \widehat{\sigma}_{p,n}^2 / N_n)$, where $\widehat{\sigma}_{p,n}^2$ is computed from a single data set. The similarities between the two calculations are easily recognizable, thus implying that the asymptotic normality method yields satisfactory proxies of the distribution of $\widehat{EAR}_{p,n}^*$.

3.2. Bandwidth selection sensitivity analysis

In the above simulation study, we used bandwidths (3.2). This is certainly not the only choice that satisfies conditions (D₁)–(D₃). Hence, we next investigate the impact of different bandwidths on the performance of $\widehat{EAR}_{p,n}$. Specifically, we set

$$\Delta_{1,n} = \Delta_{2,n} = a n^{-b/6}, \tag{3.4}$$

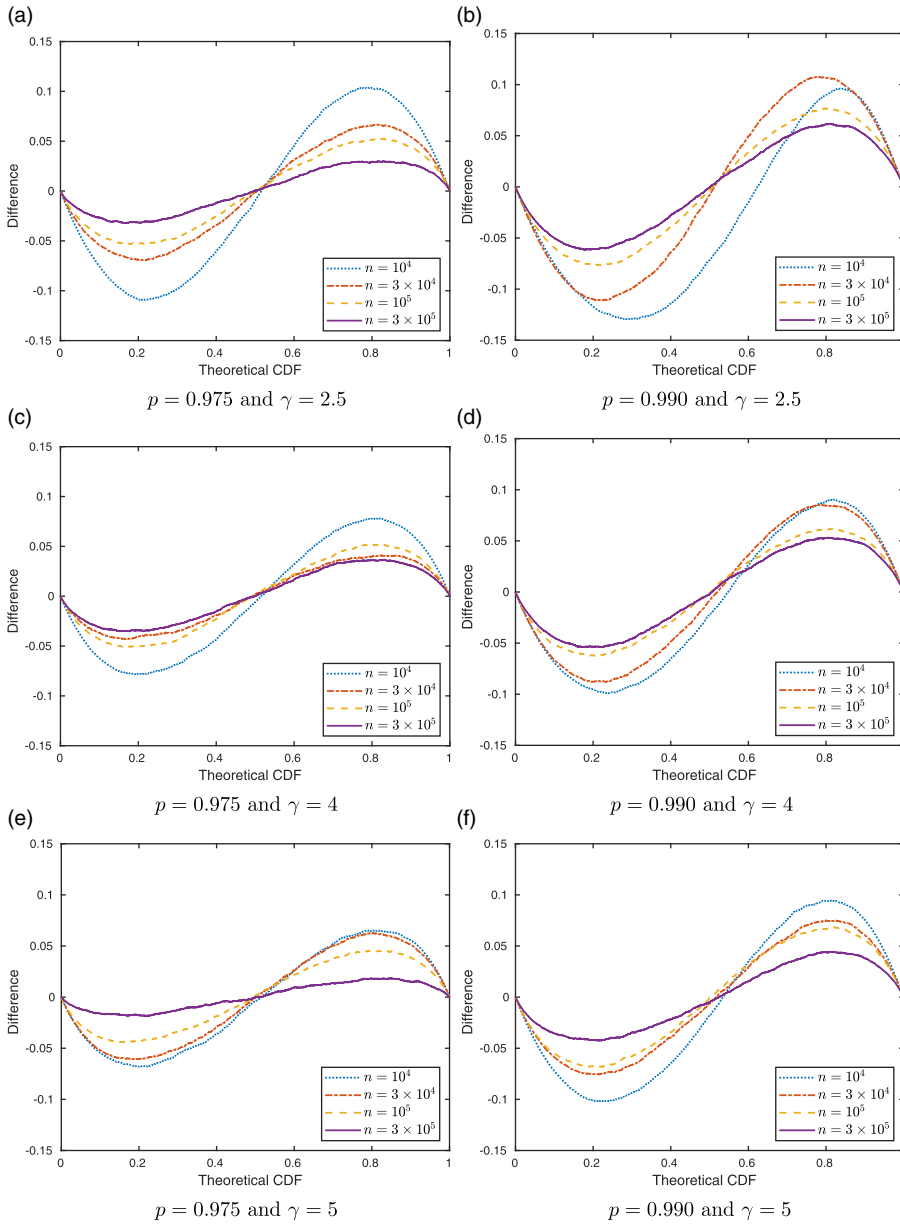


Figure 4. The difference between the cdf of $\widehat{EAR}_{p,n}^*$ and its asymptotic normality proxy based on Theorem 2.2 under various parameter and sample-size choices.

and then vary the parameters $a \in \{0.4, 0.7, 1.0, 1.3, 1.6\}$ and $b \in \{2.1, 2.4, 2.7, 3.0\}$ around the benchmark cases $a = 1$ and $b = 3$, which give $n^{-1/2}$. In view of Note 2.8, condition (D₃) holds when $b/6 \in (1/3, 1/2]$, which justifies the choices of b .

We assess the performance of $\widehat{EAR}_{p,n}$ based on the bias, standard deviation (SD), and mean absolute error (MAE) when compared with the true value of $EAR_p(X|Y)$. Tables 2 and 3 summarize the results of our sensitivity analysis when the sample size n is moderately small ($n = 10,000$) and large ($n = 300,000$). We observe from Table 2 that among all the considered scenarios, a smaller value of a leads to a better performance of $\widehat{EAR}_{p,n}$ in terms of smaller bias. However, this is at the expense of increasing uncertainty measured in terms of the SD, except in the cases of the small sample

Table 2. $\widehat{EAR}_{p,n}$ performance as measured by the absolute value of bias, SD, and MAE, with varying a and fixed $b = 3$ in the assumed form $a n^{-b/6}$ of the bandwidth.

		$p = 0.975$						$p = 0.990$					
		$n = 10,000$			$n = 300,000$			$n = 10,000$			$n = 300,000$		
		Bias	SD	MAE	Bias	SD	MAE	Bias	SD	MAE	Bias	SD	MAE
$\gamma = 2.5$	$a = 0.4$	0.8	16.0	12.7	0.06	5.6	4.5	8.9	30.1	24.9	0.3	9.2	7.3
	$a = 0.7$	2.9	13.8	11.2	0.1	4.4	3.5	34.2	29.1	37.3	1.0	7.5	6.0
	$a = 1.0$	6.4	12.9	11.5	0.2	3.8	3.1	153.6	56.1	153.6	1.9	6.7	5.5
	$a = 1.3$	11.5	12.5	13.8	0.4	3.5	2.8	113.4	49.4	113.4	3.2	6.2	5.6
	$a = 1.6$	19.0	12.4	19.5	0.5	3.2	2.6	79.9	44.2	80.0	4.9	5.9	6.2
EAR_p		301.0						481.8					
$\gamma = 4$	$a = 0.4$	0.2	6.2	5.0	0.001	2.3	1.8	2.4	10.0	8.1	0.07	3.3	2.6
	$a = 0.7$	0.9	5.2	4.2	0.03	1.8	1.4	9.2	9.1	10.6	0.2	2.6	2.1
	$a = 1.0$	2.0	4.8	4.1	0.06	1.5	1.2	35.1	11.4	35.1	0.5	2.3	1.9
	$a = 1.3$	3.6	4.6	4.7	0.1	1.4	1.1	24.0	10.2	24.0	0.9	2.1	1.8
	$a = 1.6$	5.9	4.5	6.3	0.2	1.3	1.0	14.4	9.3	14.8	1.3	2.0	1.9
EAR_p		123.7						183.6					
$\gamma = 5$	$a = 0.4$	0.1	4.2	3.4	0.009	1.6	1.3	1.4	6.4	5.2	0.05	2.2	1.7
	$a = 0.7$	0.6	3.6	2.9	0.02	1.2	1.0	5.5	5.8	6.6	0.2	1.7	1.4
	$a = 1.0$	1.3	3.3	2.8	0.05	1.1	0.8	20.3	6.8	20.3	0.3	1.5	1.2
	$a = 1.3$	2.3	3.1	3.1	0.08	1.0	0.8	13.4	6.1	13.4	0.5	1.4	1.1
	$a = 1.6$	3.7	3.0	4.0	0.1	0.9	0.7	7.4	5.6	7.8	0.9	1.3	1.2
EAR_p		83.9						122.9					

($n = 10,000$), heavy tailed distribution ($\gamma = 2.5$ and 4), and high confidence probability ($p = 0.990$). These slightly inconsistent patterns have likely been caused by the simulation noise. The MAE, which depends on the interplay between the bias and the uncertainty of the estimator, supports smaller (resp. larger) values of a when the sample size n is relatively small (resp. large).

Table 3 shows that the impact of the power coefficient b on $\widehat{EAR}_{p,n}$ is more pronounced than that of the multiplicative coefficient a . As the value of b decreases or, equivalently, as $\Delta_{1,n}$ and $\Delta_{2,n}$ become larger, the bias of the estimator increases, yet the SD decreases. When both n and p are relatively small (i.e., $n = 10,000$ and $p = 0.975$), or when both are large (i.e., $n = 300,000$ and $p = 0.990$), the MAE is in favor of the benchmark choice $b = 3.0$, thus suggesting bandwidths (3.2). Otherwise, the MAE is in favor of smaller values of b .

Overall, we have found that the performance of $\widehat{EAR}_{p,n}$ in response to varying multiplicative coefficient a is more predictable than changing the power coefficient b . Therefore, we suggest the user of our method to set $b = 3$ and tune the multiplicative coefficient a . Further, our sensitivity study has shown that the bandwidths influence the trade-off between the bias and the uncertainty associated with the estimator. A smaller value of a should be considered when bias is the major concern and a moderately large value of a when controlling the estimator's uncertainty. For balancing the bias and the uncertainty, bandwidths (3.2) seem to be a good choice. Finally, the impact of bandwidth parameters on the EAR estimator seems to be rather intricate and nonlinear. It would be interesting and important to develop a more rigorous way for identifying optimal choices of a and b .

Table 3. $\widehat{EAR}_{p,n}$ performance as measured by the absolute value of bias, SD, and MAE, with fixed $a = 1$ and varying b in the assumed form $a n^{-b/6}$ of the bandwidth.

		$p = 0.975$						$p = 0.990$					
		$n = 10,000$			$n = 300,000$			$n = 10,000$			$n = 300,000$		
		Bias	SD	MAE	Bias	SD	MAE	Bias	SD	MAE	Bias	SD	MAE
$\gamma = 2.5$	$b = 2.1$	56.5	19.3	56.5	10.0	2.3	10.0	72.9	24.5	73.5	126.1	9.3	126.1
	$b = 2.4$	106.7	24.4	106.7	2.6	2.6	3.0	4.4	33.7	24.3	28.0	5.3	28.0
	$b = 2.7$	18.4	12.4	19.0	0.9	3.1	2.5	81.9	44.5	82.1	6.8	5.7	7.4
	$b = 3.0$	6.4	12.9	11.5	0.2	3.8	3.1	153.6	56.1	153.6	1.9	6.7	5.5
EAR_p		301.0						481.8					
$\gamma = 4$	$b = 2.1$	11.9	4.6	11.9	3.1	0.8	3.1	31.5	5.6	31.5	27.5	1.9	27.5
	$b = 2.4$	28.0	5.6	28.0	0.8	1.0	1.05	7.7	7.4	8.9	7.6	1.7	7.6
	$b = 2.7$	5.7	4.5	6.1	0.2	1.2	0.98	15.0	9.3	15.3	1.9	1.9	2.2
	$b = 3.0$	2.0	4.8	4.1	0.06	1.5	1.2	35.1	11.4	35.1	0.5	2.3	1.9
EAR_p		123.7						183.6					
$\gamma = 5$	$b = 2.1$	6.5	2.9	6.5	2.0	0.6	2.0	21.8	3.5	21.8	15.6	1.2	15.6
	$b = 2.4$	17.0	3.5	17.0	0.5	0.7	0.69	6.5	4.5	6.9	4.6	1.1	4.6
	$b = 2.7$	3.6	3.0	3.9	0.2	0.8	0.67	7.8	5.7	8.2	1.2	1.2	1.4
	$b = 3.0$	1.3	3.3	2.8	0.05	1.1	0.8	20.3	6.8	20.3	0.3	1.5	1.2
EAR_p		83.9						122.9					

4. An analysis of real data

Having thus developed our understanding how the estimator works on (simulated) data, and in particular how to choose bandwidth parameters, we now study the allocated loss adjustment expenses (ALAE) considered in Frees and Valdez (1998), which have been widely used in the insurance literature for studying multivariate risk modeling and measurement. The data contain 1500 records of liability claims provided by the Insurance Service Office Inc. Each entry of the data contains the indemnity payment, denoted by L_1 in our following consideration, and also the associated ALAE costs that include the adjuster expenses, legal fees, investigation costs, etc., which we collectively denote by L_2 . Intuitively, the larger the loss amount, the higher the ALAE cost, and so there is a positive dependence between L_1 and L_2 . We are interested in estimating the risk contribution of the ALAE cost to the total claim cost among the tail risk scenarios via $EAR_p(X | Y)$ with $X = L_2$ and $Y = L_1 + L_2$.

Due to the presence of indemnity limit, some indemnity payments are censored, and hence the distribution of L_1 is not continuous everywhere. However, there is no indemnity limit applied on the ALAE costs, and so the distribution of L_2 is continuous. For illustrative purposes, Figure 5 displays the empirical cdf's of the logarithmically transformed indemnity-payments and total-cost amounts. The empirical cdf associated with L_1 has uneven jump sizes, whereas we do not see jumps of significantly different sizes in the empirical cdf of Y . This confirms our statement that the population distribution of Y is continuous, and thus we safely conclude that condition (C_1) should be met. Moreover, the presence of indemnity limit also makes us to naturally accept the finite moment conditions (C_2) and (C_4) . These conditions readily ensure the asymptotic precision of $EAR_{p,n}$ according to Theorem 2.1. To utilize the asymptotic normality result established in Theorem 2.2 for constructing confidence intervals for $EAR_p(X | Y)$, we find it appropriate to assume a model satisfying condition (C_3) with $\alpha = 1$.

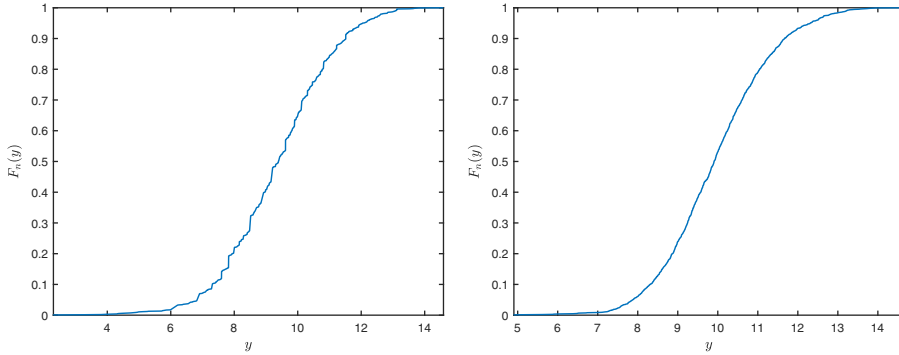


Figure 5. The empirical cdf associated with $\log(L_1)$ in the left-hand panel, and that with $\log(Y)$ where $Y = L_1 + L_2$ in the right-hand panel.

Of course, our proposed estimator can be used on data of any size and at any confidence level $p \in (0, 1)$. Since the ALAE data under investigation have a relatively small sample size, we choose to work with smaller probability levels $p = 0.8$ and $p = 0.9$ so that our derived conclusions would have higher credibility. In practice, of course, insurance companies possess much larger data, or the EAR is estimated based on synthetic data (e.g., Gabrielli and Wüthrich, 2018; Millosovich *et al.*, 2021). In the latter case, the sample sizes can be arbitrarily large, and thus the application of $\widehat{\text{EAR}}_{p,n}$ would yield EAR estimates at any probability level and at any desired precision.

Motivated by the sensitivity analysis of Section 3.2, we estimate $\text{EAR}_p(X | Y)$ by applying $\widehat{\text{EAR}}_{p,n}$ with varying multiplicative coefficients a . Our estimation results are summarized in Figure 6, which includes EAR estimates, estimation uncertainty as assessed by $\hat{\sigma}_{p,n}$ of Theorem 2.3, and 90% confidence intervals for $\text{EAR}_p(X | Y)$. In response to varying choices of a , the EAR estimates fluctuate from around 1.6×10^4 to 1.8×10^4 when $p = 0.8$, and from around 2.6×10^4 to 2.8×10^4 when $p = 0.9$. Note that as a decreases (equivalently, the estimation bandwidth becomes smaller), the standard error of $\widehat{\text{EAR}}_{p,n}$ increases. Consequently, the confidence intervals for EAR_p become wider.

The developed EAR estimation can immediately be appreciated by risk analysts. To illustrate, recall that $\text{EAR}_p(X | Y)$ is the Euler allocation when the aggregate risk is measured by $\text{VaR}_p(Y)$. Define the corresponding allocation ratio

$$r_p^{\text{EAP}}(X | Y) = \frac{\text{EAR}_p(X | Y)}{\text{VaR}_p(Y)}.$$

Based on the ALAE data, we find that $\text{VaR}_{0.8}(Y) = 6.26 \times 10^4$ and $\text{VaR}_{0.9}(Y) = 1.17 \times 10^5$. For instance, if $a = 1.0$, then $\widehat{\text{EAR}}_{p,n}$ is equal to 1.67×10^4 when $p = 0.8$, and to 2.61×10^4 when $p = 0.9$, which implies that the ALAE cost accounts for

$$r_{0.8}^{\text{EAP}}(X | Y) = 26.68\%$$

and

$$r_{0.9}^{\text{EAP}}(X | Y) = 22.31\%$$

of the aggregation risk.

The confidence intervals of $\text{EAR}_p(X | Y)$ can be used by risk analysts to understand the uncertainty associated with the EAR estimates. These intervals can also help to assess the adequacy of risk allocation that is determined by another method. For example, Chen *et al.* (in press) applied the mixed gamma distribution to estimate the TCA of the ALAE cost

$$\text{TCA}_p(X | Y) = \mathbb{E}(X | Y > \text{VaR}_p(Y))$$

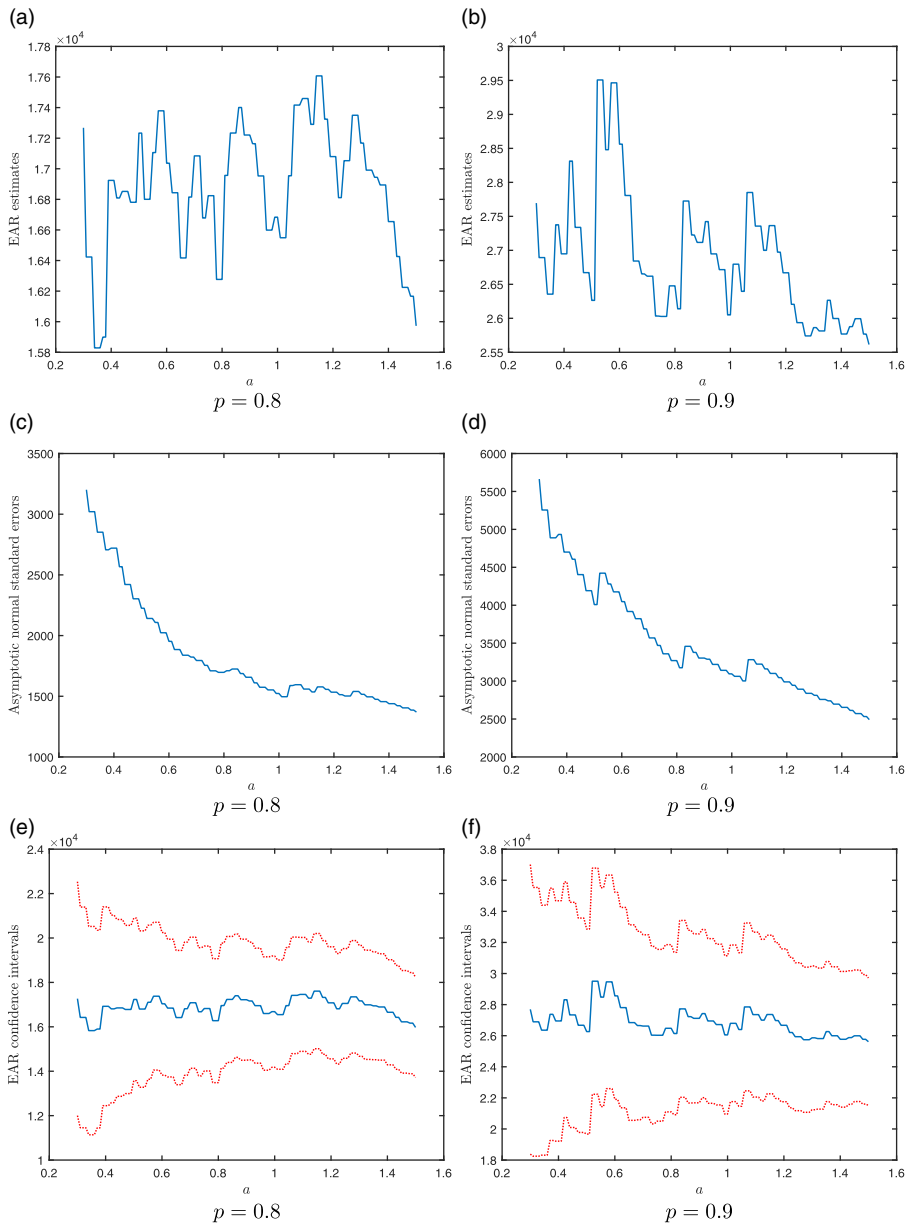


Figure 6. The ALAE data’s EAR estimates (top panels), asymptotic normal standard errors (middle panels), and the associated 90% EAR confidence intervals with lower and upper bounds depicted as dotted-red lines (bottom panels).

and the associated TCA-based allocation percentage

$$r_p^{\text{TCA}}(X | Y) = \frac{\text{TCA}_p(X | Y)}{\mathbb{E}(Y | Y > \text{VaR}_p(Y))}.$$

The authors concluded that the risk contribution of the ALAE cost is

$$r_{0.8}^{\text{TCA}}(X | Y) = 19.56\%$$

and

$$r_{0.9}^{\text{TCA}}(X | Y) = 17.86\%.$$

We can therefore conclude that the risk contribution portion determined by aforementioned TCA method is not adequate in the context of EAR allocation because the corresponding risk allocations

$$\text{VaR}_{0.8}(Y) \times r_{0.8}^{\text{TCA}}(X | Y) = 6.26 \times 10^4 \times 0.1956 = 1.22 \times 10^4$$

and

$$\text{VaR}_{0.9}(Y) \times r_{0.9}^{\text{TCA}}(X | Y) = 1.17 \times 10^5 \times 0.1786 = 2.09 \times 10^4$$

are outside of the confidence intervals of $\text{EAR}_p(X | Y)$ when $p = 0.8$ and $p = 0.9$, respectively, at least when $a = 1$.

5. Conclusion

Using the notation of concomitants, in this paper, we have defined a consistent empirical estimator of the VaR-induced EAR, established its asymptotic normality, and also proposed a consistent empirical estimator of the standard deviation. The performance of theoretical results has been illustrated using simulated and real data sets. These results facilitate statistical inference in a number of finance and insurance related contexts, such as risk-adjusted capital allocations in financial management and risk sharing in peer-to-peer insurance.

Naturally, the obtained results have generated a number of thoughts for future research, and we have already mentioned some of them. Additionally, an interesting research direction would be to develop statistical inference for the VaR-induced EAR in scenarios under dependent-data generating processes, such as time series, whose importance is seen from the study of, for example, Asimit *et al.* (2019). For this, in addition to classical methods of time series (e.g., Brockwell and Davis, 1991), the techniques developed by Sun *et al.* (2022), although in a different context, could facilitate the development of an appropriate statistical inference theory.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/asb.2023.17>.

References

- Abdikerimova, S. and Feng, F. (2022) Peer-to-peer multi-risk insurance and mutual aid. *European Journal of Operational Research*, **299**, 735–749.
- Alai, D.H., Landsman, Z. and Sherris, M. (2016) Modelling lifetime dependence for older ages using a multivariate Pareto distribution. *Insurance: Mathematics and Economics*, **70**, 272–285.
- Asimit, V., Peng, L., Wang, R. and Yu, A. (2019) An efficient approach to quantile capital allocation and sensitivity analysis. *Mathematical Finance*, **29**, 1131–1156.
- BCBS (2016) *Minimum Capital Requirements for Market Risk. January 2016*. Basel Committee on Banking Supervision. Bank for International Settlements, Basel.
- BCBS (2019). *Minimum Capital Requirements for Market Risk. February 2019*. Basel Committee on Banking Supervision. Bank for International Settlements, Basel.
- Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods*, 2nd edition. New York: Springer.
- Borovkov, A.A. (1998) *Mathematical Statistics*. London: Routledge.
- Chen, Y., Song, Q. and Su, J. (in press) On a class of multivariate mixtures of gamma distributions: Actuarial applications and estimation via stochastic gradient methods. *Variance*.
- David, H.A. and Nagaraja, H.N. (2003) *Order Statistics*, 3rd edition. New York: Wiley.
- Denuit, M. (2019) Size-biased transform and conditional mean risk sharing, with application to P2P insurance and tontines. *ASTIN Bulletin*, **49**, 591–617.
- Denuit, M., Dhaene, J. and Robert, C.Y. (2022) Risk-sharing rules and their properties, with applications to peer-to-peer insurance. *Journal of Risk and Insurance*, **89**, 615–667.
- Denuit, M. and Robert, C.Y. (2020) Large-loss behavior of conditional mean risk sharing. *ASTIN Bulletin*, **50**, 1093–1122
- EIOPA (2016) *Solvency II*. The European Insurance and Occupational Pensions Authority, Frankfurt am Main.

- ExtremeMode (2023) ExtremeMode—Handling Method for Extreme Data. The MathWorks, Inc. <https://www.mathworks.com/help/stats/boxplot.html>
- Frees, E. and Valdez, E.A. (1998) Understanding relationships using copulas. *North American Actuarial Journal*, **2**, 1–25.
- Fu, M.C., Hong, L.J. and Hu, J.-Q. (2009) Conditional Monte Carlo estimation of quantile sensitivities. *Management Science*, **55**, 2019–2027.
- Furman, E. and Zitikis, R. (2008) Weighted risk capital allocations. *Insurance: Mathematics and Economics*, **43**, 263–269.
- Furman, E. and Zitikis, R. (2009) Weighted pricing functionals with applications to insurance: An overview. *North American Actuarial Journal*, **13**, 483–496.
- Gabrielli, A. and Wüthrich, M.V. (2017) An individual claims history simulation machine. *Risks*, **6**, 1–29.
- Gourieroux, C., Laurent, J.P. and Scaillet, O. (2000) Sensitivity analysis of Values at Risk. *Journal of Empirical Finance*, **7**, 225–245.
- Gribkova, N., Su, J. and Zitikis, R. (2022a) Empirical tail conditional allocation and its consistency under minimal assumptions. *Annals of the Institute of Statistical Mathematics*, **74**, 713–735.
- Gribkova, N., Su, J. and Zitikis, R. (2022b) Inference for the tail conditional allocation: Large sample properties, insurance risk assessment, and compound sums of concomitants. *Insurance: Mathematics and Economics*, **107**, 199–222.
- Gribkova, N. and Zitikis, R. (2017) Statistical foundations for assessing the difference between the classical and weighted-Gini betas. *Mathematical Methods of Statistics*, **26**, 267–281.
- Gribkova, N. and Zitikis, R. (2019) Weighted allocations, their concomitant-based estimators, and asymptotics. *Annals of the Institute of Statistical Mathematics*, **71**, 811–835.
- Hong, L.J. (2009) Estimating quantile sensitivities. *Operations Research*, **57**, 118–130.
- IAIS (2016) *2016 Risk-based Global Insurance Capital Standard (ICS) Consultation Document*. The International Association of Insurance Supervisors, Basel.
- James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013) *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.
- Jiang, G. and Fu, M.C. (2015) Technical note – on estimating quantile sensitivities via infinitesimal perturbation analysis. *Operations Research*, **63**, 435–441.
- Koenker, R. (2005) *Quantile Regression*. Cambridge: Cambridge University Press.
- Liu, G. and Hong, L.J. (2009) Kernel estimation of quantile sensitivities. *Naval Research Logistics*, **56**, 511–525.
- McNeil, A.J., Frey, R. and Embrechts, P. (2015) *Quantitative Risk Management: Concepts, Techniques and Tools, Revised Edition*. Princeton: Princeton University Press.
- Miliosovich, P., Tsanakas, A. and Wang, R. (2021) A theory of multivariate stress testing. Available at SSRN: <https://ssrn.com/abstract=3966204>.
- Rao, C.R. and Zhao, L.C. (1995) Convergence theorems for empirical cumulative quantile regression function. *Mathematical Methods of Statistics*, **4**, 81–91.
- Sarabia, J.M., Gómez-Déniz, E., Prieto, F. and Jordá, V. (2018) Aggregation of dependent risks in mixtures of exponential distributions and extensions. *ASTIN Bulletin*, **48**, 1079–1107.
- Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*. Boca Raton: Chapman and Hall.
- Su, J. and Furman, E. (2016) A form of multivariate Pareto distribution with applications to financial risk measurement. *ASTIN Bulletin*, **47**, 331–357.
- Sun, N., Yang, C. and Zitikis, R. (2022) Detecting systematic anomalies affecting systems when inputs are stationary time series. *Applied Stochastic Models in Business and Industry*, **38**, 512–544.
- Tasche, D. (2007) *Capital Allocation to Business Units and Sub-Portfolios: the Euler Principle*. Technical Report, arXiv:0708.2542v3. <https://doi.org/10.48550/arXiv.0708.2542>
- Tasche, D. (2009) Capital allocation for credit portfolios with kernel estimators. *Quantitative Finance*, **9**, 581–595.
- Tse, S.M. (2009) On the cumulative quantile regression process. *Mathematical Methods of Statistics*, **18**, 270–279.