STATISTICALLY (OPTIMAL) ESTIMATORS OF SEMIVARIANCE: A CORRECTION OF JOSEPHY-ACZEL'S PROOF

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 17(23)

Karlheinz Fleischer, Bernhard Nietert
School of Business and Economics,
Philipps-University Marburg, Germany
e-mails: k.fleischer@wiwi.uni-marburg.de; nietert@wiwi.uni-marburg.de

ISSN 1644-6739
e-ISSN 2449-9765

DOI: 10.15611/sps.2019.17.01
JEL Classification: C49, G19

Abstract: Semivariance is an intuitive risk measure because it concentrates on the shortfall below a target and not on total variation. To successfully use semivariance in practice, however, a statistical estimator of semivariance is needed; Josephy and Aczel provide such an estimator. Unfortunately, they have not correctly proven asymptotic unbiasedness and mean squared error consistency of their estimator since their proof contains a mistake. This paper corrects the computational mistake in Josephy-Aczel's original proof and, that way, allows researchers and practitioners in the field of downside portfolio selection, hedging, downside asset pricing, risk measurement in a regulatory context, and performance measurement to work with a meaningfully specified downside measure.

Keywords: risk analysis, semivariance, statistical estimation.

1. Introduction to the problem

On the one hand, semivariance – lower partial moment 2 – is an intuitive risk measure because it concentrates on the shortfall below a target and not on total variation. Therefore it is used in downside portfolio selection (pioneered by [Markowitz 1959; Jin, Markowitz, Zhou 2006]) for an overview of and further developments in downside portfolio selections, hedging [Demirer, Lien 2003; Cotter, Hanly 2006], downside asset pricing (pioneered by [Bawa, Lindenberg 1977; Ang, Chen, Xing 2006] for an overview of and further developments in downside betas), risk measurement in a regulatory context [Brooks, Persand 2003], and performance measurement (for stocks, e.g. [Hoechner, Reichling Schulze 2017], for the effects of environmental, social, and governance issues [Hoepner et al. 2018], and for hedging [Lee, Chien 2010]. On the other hand, the formalism behind lower partial moments is identical to the computation of Conditional Value at Risk [Demirer, Lien, Shaffer 2005, p. 56] were the first to emphasize this connection – and stochastic dominance [Davidson, Duclas 2000, p. 1444].

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

To successfully use semivariance and its formal companion in practice, a statistical estimator of semivariance is needed. In this connection, two aspects must be addressed: first, it must be determined when semivariance is triggered; second, the target must be estimated, for example, if mean is used as target. In the latter case, both aspects will be interrelated since the estimator of semivariance is triggered based on the sample average instead of the unknown true mean. In that case, only the trigger problem exists, Davidson and Duclas [2000] derive a rather general estimator where observations just have to be independent, but not identically distributed, but are less successful for the interrelated case: they determine an estimator for lower partial moment 1 for identically distributed and independent observations, but not an estimator for semivariance. Unfortunately, this interrelated case is the economically relevant one. For example, Lewis [1990] recommends the mean as target for lower partial moment based performance measurement. Barrett and Donald [2003, p. 71] stress that it is important to compare more than one point when ranking alternatives based on stochastic dominance – using the mean as target might mitigate this problem.

Only the vastly underappreciated paper by Josephy and Aczel [1993] provides a statistical estimator for semivariance where target equals mean and observations are independent and identically distributed. Even though independent and identically distributed observations are rather mundane and do not seem to be representative of modern-day econometrics [Josephy, Aczel 1993] is still the most advanced paper in the field of estimating semivariance with mean as target.

Unfortunately, Josephy and Aczel [1993] have not correctly proven the asymptotic unbiasedness and mean squared error consistency of their estimator. In their proof, they mistakenly assume terms to be 0 that are in fact unequal to 0.

Given this setting, the goal of this paper is to identify and correct the mistake in Josephy-Aczel's [1993] derivation. In this connection, we deviate from their (1993) original proof of asymptotic unbiasedness and mean squared error consistency. Their computational mistake results in several missing terms which can be handled more easily if a modified, and in our opinion faster, line of reasoning is used.

We show that the terms overlooked by Josephy and Aczel [1993] approach zero as the number of observations approaches infinity. Therefore, we can prove that Josephy-Aczel's [1993] estimator is indeed asymptotically unbiased and mean squared error consistent.

Compared to the literature we can make two contributions. First and obviously, we identify and correct the mistake in [Josephy, Aczel 1993]. Second, in that way we allow researchers and practitioners in the

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

field of downside portfolio selection, hedging, downside asset pricing, risk measurement in a regulatory context, and performance measurement to work with a meaningfully specified downside measure, i.e. semivariance with mean as target, instead of making unfortunate compromises, i.e. using a constant as target when applying lower partial moment 2 or falling back to lower partial moment 1 when mean is used as target.

The remainder of the paper is organized as follows. In Section 2 the computational mistake in Josephy-Aczel's [1993] derivation is identified and in Section 3 it is corrected. The paper ends with a summary (Section 4) and a mathematical appendix.

2. Notations and identification of the computational mistake

Notations

Since an immediate comparison to [Josephy, Aczel 1993] is essential for our paper, we use their notation:

Let $x_1, \dots, x_n, \dots$ denote a sequence of independent and identically distributed random variables with a finite fourth moment.

We define for $i = 1, \dots, n$:

$$y_i := x_i - \mu, \tag{1}$$

where $\mu := E\{x_1\}$.

When referring to aspects that hold for all x_i (y_i), like the expected value or the variance, x_1 (y_1) is taken to represent all x_i (y_i) because x_i (and hence y_i) are independent and identically distributed random variables.

Given the definition of y_1, it is true

$$E\{y_1\} = 0 \qquad \sigma^2 := E\{y_1^2\} = var(y_1) = var(x_1) \tag{2}$$

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu) = \bar{x} - \mu, \tag{3}$$

with $\quad \bar{x} := \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$.

Furthermore, we define

$$m_i := 1_{(-\infty,0]}(y_i) = 1_{(-\infty,0]}(x_i - \mu) \qquad l_i := 1_{(0,\bar{y}]}(y_i) - 1_{(\bar{y},0]}(y_i), \tag{4}$$

where $1_A(z)$ denotes the value of the indicator function on a set A for variable z.

The value m_i indicates when semivariance is "triggered" and l_i captures "triggering errors". The estimator of semivariance is triggered based on the sample average $\bar{y}$ instead on the true mean μ.

Semivariance reads

$$\sigma_-^2 := E\{1_{(-\infty,0]}(x_1-\mu)\cdot(x_1-\mu)^2\} = E\{1_{(-\infty,0]}(y_1)\cdot y_1^2\} = E\{m_1\cdot y_1^2\} \tag{5}$$

which suggests as reasonable estimator t_n

$$t_n := c_n \cdot \sum_{i=1}^{n} 1_{(-\infty,0]}(x_i-\bar{x})\cdot(x_i-\bar{x})^2 = c_n \cdot \sum_{i=1}^{n} 1_{(-\infty,0]}(y_i-\bar{y})\cdot(y_i-\bar{y})^2 \tag{6}$$

where the multiplier $c_n \in \mathbb{R}$ is yet to be determined.

Using the definitions of m_i and l_i, Equation (6) simplifies to

$$t_n = c_n \cdot \sum_{i=1}^{n}[m_i + l_i]\cdot(y_i-\bar{y})^2. \tag{7}$$

Furthermore, set for $\tau = 0, 1, 2, \ldots$

$$v_\tau := E\{m_1\cdot(x_1-\mu)^\tau\} = E\{m_1\cdot y_1^\tau\}, \tag{8}$$

where it obviously holds: $v_2 = E\{m_1\cdot y_1^2\} = \sigma_-^2$ and v_0 can be interpreted as probability of $y_1 \le 0$, which corresponds to the probability of $x_1 \le \mu$.

2.1. Identification of the computational mistake in Josephy-Aczel's proof

To distinguish subsequently between "our" and "Josephy-Aczel's" formulas, we put JA in front of their formula numbers.

Josephy and Aczel claim

$$E\{m_i\cdot y_j\cdot y_k\} = \begin{cases} E\{m_i\cdot y_i^2\} = \sigma_-^2 & i=j=k \\ 0 & otherwise \end{cases}. \tag{JA14}$$

While it is true, as Josephy and Aczel state, that $E\{m_i\cdot y_j\cdot y_k\} = 0$ for $j \neq k$ due to the independence of y_j and y_k and the fact that $E\{y_j\} = E\{y_k\} = 0$, this is not true for $j = k \neq i$. For $j = k \neq i$ it holds

$$E\{m_i\cdot y_j\cdot y_k\} = E\{m_i\cdot y_j^2\} = v_0\cdot E\{y_j^2\} = v_0\cdot\sigma^2. \tag{9}$$

Term (9) is missing in Josephy-Aczel's [1993] proof of asymptotic unbiasedness and mean squared error consistency of their estimator, one mistake that leads subsequently to several wrong formulas.

Josephy-Aczel's [1993] proof of asymptotic unbiasedness is based on Formula (JA17). Since (JA17) rests upon (JA14), the term discovered in

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

Formula (9) is missing causing a follow-up error. The corrected formula (the notation "corr" is added to the numbering of the formulas) reads:

$$\mathrm{E}\begin{Bmatrix} t_n \\ c_n \end{Bmatrix} = \underbrace{\frac{(n-1)^2}{n} \cdot \sigma_-^2 + \sum_{i=1}^{n} E\{l_i \cdot (y_i - \bar{y})^2\}}_{(JA17)} + \frac{n-1}{n} \cdot \sigma^2 \cdot v_0. \quad \text{(corrJA17)}$$

Josephy-Aczel's [1993] proof of mean squared error consistency is based on their Formulas (JA20a) through (JA21f), which ultimately partially rely on (JA14). For that reason, only (JA20b), (JA21a), and (JA21b) are correct.

- Formula (corrJA20a)

(JA20a) contains a typo and reads corrected (Josephy and Aczel [1993] obtain v_2^2):

$$E\{m_i \cdot m_j \cdot y_i \cdot y_j\} = v_1^2 \text{ for } i \neq j. \quad \text{(corrJA20a)}$$

All other formulas are subject to the follow-up error as the term discovered in Formula (9) is missing. The corrected formulas read (where $i \neq j$):

- Formula (corrJA20c)

$$E\{m_i \cdot m_j \cdot y_i \cdot y_j \cdot \bar{y}^2\} = \underbrace{\frac{2}{n^2} \cdot (v_1 \cdot v_3 + v_2^2)}_{(JA20c)} + \frac{n-2}{n^2} \cdot v_1^2 \cdot \sigma^2. \quad \text{(corrJA20c)}$$

- Formula (corrJA20d)

$$E\left\{m_i \cdot m_j \cdot (y_i + y_j)^2 \cdot \bar{y}^2\right\} = \underbrace{\frac{2}{n^2} \cdot (v_0 \cdot v_4 + 4 \cdot v_1 \cdot v_3 + 3 \cdot v_2^2) +}_{(JA20d)} 2 \cdot \frac{n-2}{n^2} \cdot [v_0 \cdot v_2 \cdot \sigma^2 + v_1^2 \cdot \sigma^2]. \quad \text{(corrJA20d)}$$

– Formula (corrJA20e)

$$E\{m_i \cdot m_j \cdot (y_i + y_j) \cdot \bar{y}^3\} = \underbrace{\frac{2}{n^3} \cdot (v_0 \cdot v_4 + 4 \cdot v_1 \cdot v_3 + 3 \cdot v_2^2) +}_{(JA20e)} 2 \cdot \frac{n-2}{n^3} \cdot [3 \cdot v_0 \cdot v_2 \cdot \sigma^2 + 3 \cdot v_1^2 \cdot \sigma^2 + v_0 \cdot v_1 \cdot E\{y_1^3\}]. \quad \text{(corrJA20e)}$$

– Formula (corrJA20f)

$$E\{m_i \cdot m_j \cdot \bar{y}^4\} = \underbrace{\frac{2}{n^4} \cdot (v_0 \cdot v_4 + 4 \cdot v_1 \cdot v_3 + 3 \cdot v_2^2)}_{(JA20f)} + 2 \cdot \frac{n-2}{n^4} \cdot \left[\frac{1}{2} \cdot v_0^2 \cdot E\{y_1^4\} + 4 \cdot v_0 \cdot v_1 \cdot E\{y_1^3\} + 6 \cdot v_2 \cdot \sigma^2 + \frac{3}{2} \cdot (n-3) \cdot v_0^2 \cdot \sigma^4 + 6 \cdot v_1^2 \cdot \sigma^2\right]. \quad \text{(corrJA20f)}$$

– Formula (corrJA21c)

$$E\{m_i \cdot y_i^2 \cdot \bar{y}^2\} = \underbrace{\frac{v_4}{n^2}}_{(JA21c)} + \frac{(n-1) \cdot v_2 \cdot \sigma^2}{n^2}. \quad \text{(corrJA21c)}$$

– Formula (corrJA21d)

$$E\{m_i \cdot y_i \cdot \bar{y}^3\} = \underbrace{\frac{v_4}{n^3}}_{(JA21d)} + \frac{n-1}{n^3} \cdot [v_1 \cdot E\{y_1^3\} + 3 \cdot v_2 \cdot \sigma^2]. \quad \text{(corrJA21d)}$$

– Formula (corrJA21e)

$$E\{m_i \cdot \bar{y}^4\} = \underbrace{\frac{v_4}{n^4}}_{(JA21e)} + \frac{n-1}{n^4} \cdot [v_0 \cdot E\{y_1^4\} + 4 \cdot v_1 \cdot E\{y_1^3\} + 6 \cdot v_2 \cdot \sigma^2 + 3 \cdot (n-2) \cdot v_0 \cdot \sigma^4]. \quad \text{(corrJA21e)}$$

- Formula (corrJA21f)

$$E\{m_i \cdot (y_i - \bar{y})^2\} = \underbrace{\frac{(n-1)^2}{n^2} \cdot v_2}_{(JA21f)} + \frac{n-1}{n^2} \cdot v_0 \cdot \sigma^2. \quad \text{(corrJA21f)}$$

One exemplary proof, namely the derivation of Formula (corrJA20e), can be found in Appendix 1. The other corrected formulas can be derived in a similar way.

3. Correction of the proof

To compute asymptotic unbiasedness and mean squared error consistency, we deviate from Josephy-Aczel's [1993] original proof. The missing term (9) can be handled easier if a modified and in our opinion faster line of reasoning is used.

3.1. Useful relations, lemmas, and a corollary

This subsection contains one relation, two lemmas, and one corollary that will help simplify the proofs in Subsections 3.2 and 3.3.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

Relation

For $i = 1, \cdots, n$ it holds

$$m_i^2 = m_i \text{ and } (m_i + l_i)^2 = m_i + l_i \tag{10}$$

since m_i and $m_i + l_i = 1_{(-\infty,\bar{y}]}(y_i)$ are indicator functions.

Lemma 1

For $k \in \mathbb{N}$ and $i = 1, \cdots, n$ it holds

$$\begin{gathered}|l_i| \cdot |y_i - \bar{y}|^k = \left|1_{(0,\bar{y}]}(y_i) - 1_{(\bar{y},0]}(y_i)\right| \cdot |y_i - \bar{y}|^k \leq \\ |\bar{y}|^k = |\bar{x} - \mu|^k.\end{gathered} \tag{11}$$

For the proof of (11) observe: $|l_i| = 1$ if y_i lies between 0 and $\bar{y}$ irrespective of which of either value is less. In this case, the distance between y_i and $\bar{y}$ is not greater than $|0 - \bar{y}| = |\bar{y}|$ since $|0 - \bar{y}|$ is the maximum possible distance. If otherwise y_i does not lie between 0 and $\bar{y}$, $1_{(0,\bar{y}]}(y_i) - 1_{(\bar{y},0]}(y_i)$ will be zero and so will $|l_i|$.

Lemma 2 (see Appendix 2.1 for a proof)

Let $y_1, \dots, y_n, \dots$ denote a sequence of independent and identically distributed random variables with existing mean $E\{y_1\} = 0$ and with finite k^{th} moment $E\{{y_1}^k\}$ for some $k \in \mathbb{N}$.

Then for all $\kappa = 1, .., k$ the expected value $E\{\bar{y}^\kappa\}$ exists and it holds

$$E\{\bar{y}^\kappa\} \to 0 \text{ for } n \to \infty. \tag{12}$$

Corollary (see Appendix 2.2 for a proof)

Let $y_1, \dots, y_n, \dots$ denote a sequence of independent and identically distributed random variables with existing mean $E\{y_1\} = 0$. If for $k \in \mathbb{N}$ the expected value $E\{{y_1}^k\}$ exists for k even and $E\{{y_1}^{k+1}\}$ exists for k odd, $E\{|\bar{y}|^k\}$ exists as well and it holds

$$E\{|\bar{y}^k|\} \to 0 \text{ for } n \to \infty. \tag{13}$$

Note that the existence of $E\{{y_1}^{k+1}\}$ for k odd is required only for the proof of convergence, but not for the proof of existence of $E\{{y_1}^k\}$.

3.2. Proof of asymptotic unbiasedness

$$E\{t_n\} = c_n \cdot E\left\{\sum_{i=1}^{n} [m_i + l_i] \cdot (y_i - \bar{y})^2\right\}$$

can be re-written as

$$E\{t_n\} = c_n \cdot E\{\sum_{i=1}^{n}[m_i \cdot y_i^2 - 2 \cdot m_i \cdot y_i \cdot \bar{y} + m_i \cdot \bar{y}^2 + l_i \cdot (y_i - \bar{y})^2]\} = n \cdot c_n \cdot \left[\sigma_-^2 - 2 \cdot \underbrace{E\{m_1 \cdot y_1 \cdot \bar{y}\}}_{=:A_n} + \underbrace{E\{m_1 \cdot \bar{y}^2\}}_{=:B_n} + \underbrace{E\{l_1 \cdot (y_1 - \bar{y})^2\}}_{=:C_n}\right],$$

because $y_1, \dots, y_n$ are independent and identically distributed and $E\{m_1 \cdot y_1^2\} \underbrace{=}_{(5)} \sigma_-^2$.

To prove asymptotic unbiasedness, we will show that A_n, B_n, and C_n converge to 0 for $n \to \infty$. Since null sequences are considered, it suffices to show that $|A_n|$, $|B_n| = B_n$, and $|C_n|$ converge to 0 for $n \to \infty$.

Then asymptotic unbiasedness will hold if c_n is chosen in a way such that $n \cdot c_n$ converges to 1 for $n \to \infty$:

- A_n

$$0 \leq |A_n| = |E\{m_1 \cdot y_1 \cdot \bar{y}\}| = \left|E\left\{\frac{1}{n} \cdot m_1 \cdot y_1 \cdot \left(y_1 + \sum_{i=2}^{n} y_i\right)\right\}\right| = \frac{1}{n} \cdot \left|E\{m_1 \cdot y_1^2\} + \sum_{i=2}^{n} E\{m_1 \cdot y_1 \cdot y_i\}\right| \underbrace{=}_{(2)} \frac{1}{n} \cdot E\{m_1 \cdot y_1^2\} \underbrace{=}_{(5)} \frac{\sigma_-^2}{n} \to 0$$

for $n \to \infty$.

- B_n

$$0 \leq |B_n| = B_n = E\{m_1 \cdot \bar{y}^2\} \leq E\{\bar{y}^2\} \underbrace{\to 0}_{Lemma\ 2} \text{ for } n \to \infty.$$

- C_n

$$0 \leq |C_n| = |E\{l_1 \cdot (y_1 - \bar{y})^2\}| \underbrace{\leq}_{(\#)} E\{|l_1| \cdot |y_1 - \bar{y}|^2\} \underbrace{\leq}_{Lemma\ 1} E\{|\bar{y}|^2\} \underbrace{\to 0}_{Corollary} \text{ for } n \to \infty$$

Inequality (#) rests upon the well-known relation $|E\{X\}| \leq E\{|X|\}$.

- $n \cdot c_n$

Asymptotic unbiasedness is obtained if c_n is selected in a way such that $n \cdot c_n$ converges to 1 for $n \to \infty$. This means, a multiplier $c_n = \frac{1}{n} \cdot a_n$ can be chosen for any a_n with $a_n \to 1$ for $n \to \infty$. Special cases for $c_n(a_n)$ include $c_n = \frac{1}{n}$ $(a_n = 1)$, $c_n = \frac{1}{n-1}$ $(a_n = \frac{n}{n-1})$, Josephy-Aczel's [1993]

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

recommendation $c_n = \frac{n}{(n-1)^2}$ $(a_n = \frac{n^2}{(n-1)^2})$ or $c_n = \frac{1}{n-k}$ for any $k \in \mathbb{Z}$ and $n > k$ $(a_n = \frac{n}{n-k})$.
Observe that $c_n \to 0$ for $n \to \infty$ if $n \cdot c_n \to 1$.

3.3. Proof of mean squared error consistency

Since the fourth moment $E\{y_1^4\}$ was assumed to exist, the mean squared error of t_n exists and can be written:

$$mse(t_n) = var(t_n) + \left(bias(t_n)\right)^2 = E\{t_n{}^2\} - (E\{t_n\})^2 + \left(bias(t_n)\right)^2.$$

However, it is already known that the estimator is asymptotically unbiased so that it holds

$$E\{t_n\} \to \sigma_-^2 \text{ for } n \to \infty$$

$$(E\{t_n\})^2 \to \sigma_-^4 \text{ for } n \to \infty$$

$$bias(t_n) \to 0 \text{ for } n \to \infty$$

$$\left(bias(t_n)\right)^2 \to 0 \text{ for } n \to \infty$$

hence it remains to show that $E\{t_n{}^2\} \to \sigma_-^4$ for $n \to \infty$.

To that end, we consider

$$E\{t_n{}^2\} = E\left\{\left(c_n \cdot \sum_{i=1}^{n} (m_i + l_i) \cdot (y_i - \bar{y})^2\right)^2\right\} =$$

$$c_n{}^2 \cdot E\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} (m_i + l_i) \cdot (m_j + l_j) \cdot (y_i - \bar{y})^2 \cdot \left(y_j - \bar{y}\right)^2\right\} =$$

$$c_n{}^2 \cdot E\left\{\sum_{i=1}^{n} (m_i + l_i)^2 \cdot (y_i - \bar{y})^4\right\} +$$

$$c_n{}^2 \cdot E\left\{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} (m_i + l_i) \cdot (m_j + l_j) \cdot (y_i - \bar{y})^2 \cdot \left(y_j - \bar{y}\right)^2\right\} \underbrace{=}_{iid,(10)} \tag{14}$$

$$n \cdot c_n{}^2 \cdot \underbrace{E\{(m_1 + l_1) \cdot (y_1 - \bar{y})^4\}}_{=:D_n} +$$

$$n \cdot (n-1) \cdot c_n{}^2 \cdot \underbrace{E\{(m_1 + l_1) \cdot (m_2 + l_2) \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2\}}_{=:F_n}.$$

It is shown in Appendix 3.1 that $n \cdot {c_n}^2 \cdot D_n$ converges to 0 for $n \to \infty$ and $n \cdot (n-1) \cdot {c_n}^2 \cdot F_n$ converges to σ_-^4 (see Appendix 3.2) why

$$E\{{t_n}^2\} \to \sigma_-^4 \text{ for } n \to \infty$$

holds and mean squared error consistency of the estimator t_n is proven.

4. Summary

The starting point for our paper is the observation that semivariance is applied in downside portfolio selection, hedging, downside asset pricing, risk measurement in a regulatory context, and performance measurement. To successfully use semivariance in practice, a statistical estimator of semivariance is needed.

If mean is used as target for semivariance, only the vastly underappreciated paper by Josephy and Aczel [1993] provides a statistical estimator for semivariance. Unfortunately, Josephy and Aczel [1993] have not correctly proven asymptotic unbiasedness and mean squared error consistency of their estimator. In their proof, they mistakenly assume terms to be 0 that are in fact unequal to 0.

This paper corrects the computational mistake in Josephy and Aczel's [1993 original proof of asymptotic unbiasedness and mean squared error consistency. It shows that the terms overlooked by Josephy and Aczel [1993] approach zero as the number of observations approaches infinity. Therefore, we prove that Josephy and Aczel's [1993] estimator is indeed asymptotically unbiased and mean squared error consistent.

In that way, the paper allows researchers and practitioners in the field of downside portfolio selection, hedging, downside asset pricing, risk measurement in a regulatory context, and performance measurement to work with a meaningfully specified downside measure, i.e. semivariance with mean as target, instead of making unfortunate compromises, i.e. using a constant as target when applying lower partial moment 2 or falling back to lower partial moment 1 when mean is used as target.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

Appendix 1. Proof of formula (corrJA20e)

$$E\{m_i \cdot m_j \cdot (y_i + y_j) \cdot \bar{y}^3\} = E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\} + E\{m_i \cdot m_j \cdot y_j \cdot \bar{y}^3\} = 2 \cdot E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\}$$

Since

$$\bar{y}^3 = \frac{1}{n^3} \cdot \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{t=1}^{n} y_r \cdot y_s \cdot y_t =$$

$$\frac{1}{n^3} \cdot \left(\sum_{r=1}^{n} y_r^3 + 3 \cdot \sum_{r=1}^{n}\sum_{\substack{s=1 \\ s\neq r}}^{n} y_r^2 \cdot y_s + \sum_{r=1}^{n}\sum_{\substack{s=1 \\ s\neq r}}^{n}\sum_{\substack{t=1 \\ t\neq r,s}}^{n} y_r \cdot y_s \cdot y_t \right)$$

it is obtained for

$$E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\} = \frac{1}{n^3} \cdot$$

$$E\left\{ m_i \cdot m_j \cdot y_i \cdot \sum_{r=1}^{n} y_r^3 + m_i \cdot m_j \cdot y_i \cdot 3 \cdot \sum_{r=1}^{n}\sum_{\substack{s=1 \\ s\neq r}}^{n} y_r^2 \cdot y_s + m_i \cdot m_j \cdot y_i \right.$$

$$\left. \cdot \sum_{r=1}^{n}\sum_{\substack{s=1 \\ s\neq r}}^{n}\sum_{\substack{t=1 \\ t\neq r,s}}^{n} y_r \cdot y_s \cdot y_t \right\} =$$

$$\frac{1}{n^3} \cdot \left[E\{m_i \cdot m_j \cdot y_i^4\} + E\{m_i \cdot m_j \cdot y_i \cdot y_j^3\} + E\left\{ m_i \cdot m_j \cdot y_i \cdot \sum_{\substack{r=1 \\ r\neq i,j}}^{n} y_r^3 \right\} + \right.$$

$$3 \cdot E\{m_i \cdot m_j \cdot y_i^3 \cdot y_j\} + 3 \cdot E\left\{ m_i \cdot m_j \cdot y_i^3 \cdot \sum_{\substack{s=1 \\ s\neq i,j}}^{n} y_s \right\} +$$

$$3 \cdot E\{m_i \cdot m_j \cdot y_i^2 \cdot y_j^2\} +$$

$$3 \cdot E\left\{ m_i \cdot m_j \cdot y_i \cdot y_j^2 \cdot \sum_{\substack{s=1 \\ s\neq i,j}}^{n} y_s \right\} + 3 \cdot E\left\{ m_i \cdot m_j \cdot y_i^2 \cdot \sum_{\substack{r=1 \\ r\neq i,j}}^{n} y_r^2 \right\} +$$

$$3 \cdot E\left\{ m_i \cdot m_j \cdot y_i \cdot y_j \cdot \sum_{\substack{r=1 \\ r\neq i,j}}^{n} y_r^2 \right\} + 3 \cdot E\left\{ m_i \cdot m_j \cdot y_i \cdot \sum_{\substack{r=1 \\ r\neq i,j}}^{n}\sum_{\substack{s=1 \\ s\neq r,i,j}}^{n} y_r^2 \cdot y_s \right\} +$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 17(23)

$$E\left\{m_i \cdot m_j \cdot y_i \cdot \sum_{r=1}^{n} \sum_{\substack{s=1\\ s\neq r}}^{n} \sum_{\substack{t=1\\ t\neq r,s}}^{n} y_r \cdot y_s \cdot y_t\right\}\Bigg]$$

$$\underbrace{=}_{independence} \frac{1}{n^3} \cdot \left[E\{m_i \cdot y_i^4\} \cdot E\{m_j\} + E\{m_i \cdot y_i\} \cdot E\{m_j \cdot y_j^3\} + \right.$$

$$E\{m_i \cdot y_i\} \cdot E\{m_j\} \cdot E \sum_{\substack{r=1\\ r\neq i,j}}^{n} \{y_r^3\} +$$

$$3 \cdot E\{m_i \cdot y_i^3\} \cdot E\{m_j \cdot y_j\} + 3 \cdot E\{m_i \cdot y_i^3\} \cdot E\{m_j\} \cdot \sum_{\substack{s=1\\ s\neq i,j}}^{n} E\{y_s\} +$$

$$3 \cdot E\{m_i \cdot y_i^2\} \cdot E\{m_j \cdot y_j^2\} +$$

$$3 \cdot E\{m_i \cdot y_i\} \cdot E\{m_j \cdot y_j^2\} \cdot \sum_{\substack{s=1\\ s\neq i,j}}^{n} E\{y_s\} + 3 \cdot E\{m_i \cdot y_i^2\} \cdot E\{m_j\} \cdot \sum_{\substack{r=1\\ r\neq i,j}}^{n} E\{y_r^2\}$$

$$+ 3 \cdot E\{m_i \cdot y_i\} \cdot E\{m_j \cdot y_j\} \cdot \sum_{\substack{r=1\\ r\neq i,j}}^{n} E\{y_r^2\}$$

$$+3 \cdot E\{m_i \cdot y_i\} \cdot E\{m_j\} \cdot \sum_{\substack{r=1\\ r\neq i,j}}^{n} E\{y_r^2\} \cdot \sum_{\substack{s=1\\ s\neq r,i,j}}^{n} E\{y_s\}$$

$$+ \sum_{r=1}^{n} \sum_{\substack{s=1\\ s\neq r}}^{n} \sum_{\substack{t=1\\ t\neq r,s}}^{n} E\{m_i \cdot m_j \cdot y_i \cdot y_r \cdot y_s \cdot y_t\}.$$

Given the definition of v_τ in (8), the fact that $y_1, \dots, y_n, \dots$ are independent and identically distributed , and $E\{y_1\} = 0$ (see (2)), it is gained for

$$E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\} = \frac{1}{n^3} \cdot [v_4 \cdot v_0 + v_1 \cdot v_3 + v_1 \cdot v_0 \cdot (n-2) \cdot E\{y_1^3\} +$$
$$3 \cdot v_3 \cdot v_1 + 3 \cdot v_3 \cdot v_0 \cdot 0 +$$
$$3 \cdot v_2 \cdot v_2 + 3 \cdot v_1 \cdot v_2 \cdot 0 + 3 \cdot v_2 \cdot v_0 \cdot (n-2) \cdot \sigma^2 +$$
$$3 \cdot v_1^2 \cdot (n-2) \cdot \sigma^2 + 3 \cdot v_1 \cdot v_0 \cdot (n-2) \cdot \sigma^2 \cdot 0 + 0].$$

To understand the zero in the last term of the above formula, note that by construction, r, s, and t never possess identical values. Moreover, at least one of these subscripts is in addition different from i and j. Hence, $E\{m_i \cdot m_j \cdot y_i \cdot y_r \cdot y_s \cdot y_t\}$ is a linear function of $E\{y_1\}$.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

This finally yields

$$E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\} = \frac{1}{n^3} \cdot [v_4 \cdot v_0 + 4 \cdot v_1 \cdot v_3 + 3 \cdot v_2^2]$$
$$\frac{n-2}{n^3} \cdot [3 \cdot v_2 \cdot v_0 \cdot \sigma^2 + 3 \cdot v_1^2 \cdot \sigma^2 + v_1 \cdot v_0 \cdot E\{y_1^3\}].$$

Note that $E\{m_i \cdot m_j \cdot (y_i + y_j) \cdot \bar{y}^3\} = 2 \cdot E\{m_i \cdot m_j \cdot y_i \cdot \bar{y}^3\}$ to see the equivalence to (corrJA20e).

Appendix 2. Proof of lemma 2 and the corollary

Appendix 2.1. Proof of lemma 2

Lemma 2

Let $y_1, \ldots, y_n, \ldots$ denote a sequence of independent and identically distributed random variables with existing mean $E\{y_1\} = 0$ and with finite k^{th} moment $E\{y_1{}^k\}$ for some $k \in \mathbb{N}$.

Then for all $\kappa = 1, \ldots, k$ the expected value $E\{\bar{y}^\kappa\}$ exists and it holds

$$E\{\bar{y}^\kappa\} \to 0 \text{ for } n \to \infty. \qquad (12)$$

Proof

<u>First step</u>: existence of $E\{\bar{y}^\kappa\}$

Note that an existing k^{th} moment $E\{y_1{}^k\}$ implies that the 1st, 2nd, …, $k-1^{\text{th}}$ moments exist as well. Using the multinomial theorem (see e.g. [Kotz, Johnson 1985]), $\bar{y}^\kappa$ can be written as

$$\bar{y}^\kappa = \left(\frac{y_1 + \cdots + y_n}{n}\right)^\kappa = \frac{1}{n^\kappa} \cdot \sum_{k_1+k_2+..+k_n=\kappa} \binom{\kappa}{k_1, k_2, \ldots, k_n} \prod_{i=1}^{n} y_i^{k_i},$$

where $\binom{\kappa}{k_1, k_2, \ldots, k_n} = \frac{\kappa!}{\prod_{i=1}^n k_i!}$ denotes the multinomial coefficient.

Since all expected values $E\{y_i^{k_i}\}$ on the right hand side exist, the expected value of $\bar{y}^\kappa$exists as well.

<u>Second step</u>: proving (12) by induction

<u>Base case</u>: $\kappa = 1$

$$E\{\bar{y}\} = \frac{1}{n} \cdot E\{\sum_{i=1}^n y_i\} = \frac{1}{n} \cdot \sum_{i=1}^n E\{y_i\} \underbrace{=}_{iid} \frac{1}{n} \cdot n \cdot 0 = 0 \to 0 \text{ for } n \to \infty$$

<u>Inductive step</u>: We assume $E\{\bar{y}^\vartheta\} \to 0$ for $n \to \infty$ for all $\vartheta = 1, \ldots, \kappa - 1 < k$. We have to show that

$$E\{\bar{y}^\kappa\} \to 0 \text{ for } n \to \infty$$

is also true.

$E\{\bar{y}^{\kappa}\}$ reads

$$E\{\bar{y}^{\kappa}\} = \frac{1}{n^{\kappa}} \cdot E\left\{\left(\sum_{i=1}^{n} y_i\right) \cdot \left(\sum_{j=1}^{n} y_j\right)^{\kappa-1}\right\} \underbrace{=}_{independence}$$

$$\frac{1}{n^{\kappa}} \cdot \sum_{i=1}^{n} E\left\{y_i \cdot \left(\sum_{j=1}^{n} y_j\right)^{\kappa-1}\right\} = \frac{1}{n^{\kappa}} \cdot \sum_{i=1}^{n} E\left\{y_i \cdot \left(y_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} y_j\right)^{\kappa-1}\right\}$$

$$\underbrace{=}_{binomial\ theorem} \frac{1}{n^{\kappa}} \cdot \sum_{i=1}^{n} \left(\sum_{l=0}^{\kappa-1} \binom{\kappa-1}{l} \cdot E\left\{y_i^{l+1} \cdot \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} y_j\right)^{\kappa-1-l}\right\}\right)$$

$$\underbrace{=}_{independence} \frac{1}{n^{\kappa}} \cdot n \cdot \sum_{l=0}^{\kappa-1} \binom{\kappa-1}{l} \cdot E\{y_1^{l+1}\} \cdot E\left\{\left(\sum_{\substack{j=1 \\ j \neq 1}}^{n} y_j\right)^{\kappa-1-l}\right\}$$

$$\underbrace{=}_{E\{y_1\}=0} \frac{(n-1)^{\kappa-1}}{n^{\kappa-1}}$$

$$\cdot \sum_{l=1}^{\kappa-1} \binom{\kappa-1}{l} \cdot E\{y_1^{l+1}\} \cdot \frac{1}{(n-1)^{l}} \cdot E\left\{\left(\frac{1}{n-1} \cdot \sum_{j=2}^{n} y_j\right)^{\kappa-1-l}\right\}$$

It holds for $n \to \infty$: $\frac{(n-1)^{\kappa-1}}{n^{\kappa-1}} = \left(1 - \frac{1}{n}\right)^{\kappa-1} \to 1$, $E\{y_1^{l+1}\}$ is by assumption finite, and $\frac{1}{(n-1)^{l}} \to 0$ for $l \geq 1$. Moreover, $E\left\{\left(\frac{1}{n-1} \cdot \sum_{j=2}^{n} y_j\right)^{\kappa-1-l}\right\}$ exhibits the same convergence behavior as $E\{\bar{y}^{\kappa-1-l}\} = E\left\{\left(\frac{1}{n} \cdot \sum_{j=1}^{n} y_j\right)^{\kappa-1-l}\right\}$. The term $E\left\{\left(\frac{1}{n} \cdot \sum_{j=1}^{n} y_j\right)^{\kappa-1-l}\right\}$, however, converges to 0 for $n \to \infty$ and $l < \kappa - 1$ due to the inductive assumption; for $l = \kappa - 1$ this term reads $E\left\{\left(\frac{1}{n} \cdot \sum_{j=1}^{n} y_j\right)^{0}\right\} = 1$. Putting all the above statements together, it leads to: $E\{\bar{y}^{\kappa}\}$ converges to 0 for $n \to \infty$.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

Appendix 2.2. Proof of the corollary

Corollary

Let $y_1, \dots, y_n, \dots$ denote a sequence of independent and identically distributed random variables with existing mean $E\{y_1\} = 0$. If for $k \in \mathbb{N}$ the expected value $E\{{y_1}^k\}$ exists for k even and $E\{{y_1}^{k+1}\}$ exists for k odd, $E\{|\bar{y}|^k\}$ exists as well and it holds

$$E\{|\bar{y}^k|\} \to 0 \text{ for } n \to \infty. \tag{13}$$

Proof

<u>First case:</u> k is even, i.e. $k = 2 \cdot \varphi$ for some $\varphi \in \mathbb{N}$.

Then it holds

$$E\{|\bar{y}^k|\} = E\{\bar{y}^k\} \underbrace{\to}_{Lemma\ 2} 0 \text{ for } n \to \infty.$$

<u>Second case:</u> k is odd, i.e. $k = 2 \cdot \varphi + 1 = \varphi + \varphi + 1$ for some $\varphi \in \mathbb{N}_0$.

Then it is true for $k = 1$ (which means $\varphi = 0$)

$$0 \le E\{1 \cdot |\bar{y}^1|\} \underbrace{\le}_{\substack{Cauchy-Schwarz \\ inequality}} \sqrt{E\{1^2\} \cdot E\{\bar{y}^2\}} = \sqrt{\underbrace{E\{\bar{y}^2\}}_{\substack{\to 0 \\ Lemma\ 2}}} \to 0 \text{ for } n \to \infty$$

For $k > 1$, i.e. $\varphi \in \mathbb{N}$, it is obtained

$$0 \le E\{|\bar{y}^k|\} = E\{|\bar{y}|^\varphi \cdot |\bar{y}|^{\varphi+1}\} \underbrace{\le}_{\substack{Cauchy-Schwarz \\ inequality}} \sqrt{E\{\bar{y}^{2\cdot\varphi}\} \cdot E\{\bar{y}^{2\cdot\varphi+2}\}}$$

$$= \sqrt{\underbrace{E\{\bar{y}^{k-1}\}}_{\substack{\to 0 \\ Lemma\ 2}} \cdot \underbrace{E\{\bar{y}^{k+1}\}}_{\substack{\to 0 \\ Lemma\ 2}}} \to 0 \text{ for } n \to \infty.$$

Appendix 3. Proofs in the context of mean squared error consistency

Appendix 3.1. Analysis of $\boldsymbol{n \cdot {c_n}^2 \cdot D_n}$ from Formula (14)

$$D_n = E\{(m_1 + l_1) \cdot (y_1 - \bar{y})^4\} = E\{m_1 \cdot (y_1 - \bar{y})^4\} + E\{l_1 \cdot (y_1 - \bar{y})^4\}.$$

Applying the binomial theorem to $(y_1 - \bar{y})^4$ leads to

$$D_n = \sum_{i=0}^{4} \binom{4}{i} \cdot \underbrace{E\{m_1 \cdot y_1^i \cdot \bar{y}^{4-i}\}}_{=:G_{i,n}} + \underbrace{E\{l_1 \cdot (y_1 - \bar{y})^4\}}_{=:H_n}$$

Terms $G_{i,n}$

- $G_{4,n}$:

$$0 \leq |G_{4,n}| = G_{4,n} = E\{m_1 \cdot y_1^4\} \leq E\{y_1^4\}$$

Hence, $G_{4,n}$ is bounded since $E\{y_1^4\}$ is, by assumption, finite.

- $G_{0,n}$:

$$0 \leq |G_{0,n}| = G_{0,n} = E\{m_1 \cdot \bar{y}^4\} \leq E\{\bar{y}^4\} \underbrace{\to}_{Lemma\ 2} 0 \text{ for } n \to \infty$$

- $G_{2,n}$:

$$0 \leq |G_{2,n}| = G_{2,n} =$$

$$E\{m_1 \cdot y_1^2 \cdot \bar{y}^2\} \underbrace{\leq}_{\substack{Cauchy-Schwarz \\ inequality}} \sqrt{\underbrace{E\{m_1^2 \cdot y_1^4\}}_{=E\{m_1 \cdot y_1^4\} = G_{4,n}} \cdot \underbrace{E\{\bar{y}^4\}}_{\substack{\to 0 \\ Lemma\ 2}}} \to 0$$

for $n \to \infty$

- $G_{1,n}$:

$$0 \leq |G_{1,n}| =$$

$$|E\{m_1 \cdot y_1 \cdot \bar{y}^3\}| \underbrace{\leq}_{\substack{Cauchy-Schwarz \\ inequality}} \sqrt{\underbrace{E\{m_1^2 \cdot y_1^2 \cdot \bar{y}^2\}}_{=E\{m_1 \cdot y_1^2 \cdot \bar{y}^2\} = G_{2,n} \to 0} \cdot \underbrace{E\{\bar{y}^4\}}_{\substack{\to 0 \\ Lemma\ 2}}} \to 0$$

for $n \to \infty$

- $G_{3,n}$:

$$0 \leq |G_{3,n}| = |E\{m_1 \cdot y_1^3 \cdot \bar{y}\}| = \frac{1}{n} \cdot \left| E\{m_1 \cdot y_1^4\} + \sum_{i=2}^{n} E\{m_1 \cdot y_1^3 \cdot y_i\} \right|$$

$$\underbrace{=}_{(2)} \frac{1}{n} \cdot E\{m_1 \cdot y_1^4\} = \frac{v_4}{n} \to 0 \text{ for } n \to \infty$$

Term H_n

For term H_n it is obtained:

$0 \leq |H_n| = |E\{l_1 \cdot (y_1 - \bar{y})^4\}| \underbrace{\leq}_{(*)} E\{|l_1| \cdot |y_1 - \bar{y}|^4\} \underbrace{\leq}_{Lemma\ 1} \mathrm{E}\{|\bar{y}|^4\} \underbrace{\to 0}_{Corollary}$ for $n \to \infty$.

Inequality (*) rests upon the well-known relation $|E\{X\}| \leq E\{|X|\}$.

Term $n \cdot {c_n}^2$

Since $n \cdot c_n$ converges to 1, $n \cdot {c_n}^2$ converges to 0 for $n \to \infty$.

Summary

Putting together all the intermediate results, $n \cdot {c_n}^2 \cdot D_n$ converges to 0 for $n \to \infty$.

Appendix 3.2. Analysis of $\boldsymbol{n \cdot (n-1) \cdot {c_n}^2 \cdot F_n}$ from Formula (14)

$$F_n = E\{(m_1 + l_1) \cdot (m_2 + l_2) \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2\}$$

$$= \underbrace{E\{m_1 \cdot m_2 \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2\}}_{=:F_{1,n}} + 2$$

$$\cdot \underbrace{E\{m_1 \cdot l_2 \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2\}}_{=:F_{2,n}} + \underbrace{E\{l_1 \cdot (y_1 - \bar{y})^2 \cdot l_2 \cdot (y_2 - \bar{y})^2\}}_{=:F_{3,n}}$$

Term $F_{2,n}$

It holds

$$0 \leq |F_{2,n}| = |E\{m_1 \cdot l_2 \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2\}|$$

$$\underbrace{\leq}_{(*)} E\{|m_1 \cdot l_2 \cdot (y_1 - \bar{y})^2 \cdot (y_2 - \bar{y})^2|\} \underbrace{=}_{\substack{nonnegativity\\ of\ all\\ components}} E\{m_1 \cdot (y_1 - \bar{y})^2$$

$$\cdot |l_2| \cdot |(y_2 - \bar{y})^2|\}$$

$$\underbrace{\leq}_{Lemma\ 1} E\{m_1 \cdot (y_1 - \bar{y})^2 \cdot |\bar{y}|^2\} = E\{m_1 \cdot (y_1 - \bar{y})^2 \cdot \bar{y}^2\}$$

$$= \underbrace{E\{m_1 \cdot {y_1}^2 \cdot \bar{y}^2\}}_{=G_{2,n}} - 2 \cdot \underbrace{E\{m_1 \cdot y_1 \cdot \bar{y}^3\}}_{=G_{1,n}} + \underbrace{E\{m_1 \cdot \bar{y}^4\}}_{=G_{0,n}}.$$

Inequality (*) rests upon the well-known relation $|E\{X\}| \leq E\{|X|\}$.
Since $G_{0,n} \to 0, G_{1,n} \to 0,$ and $G_{2,n} \to 0$ for $n \to \infty$, $|F_{2,n}|$ and, thus, $F_{2,n}$ converges to 0 for $n \to \infty$.

Term $F_{3,n}$

$$0 \leq |F_{3,n}| = |E\{l_1 \cdot (y_1 - \bar{y})^2 \cdot l_2 \cdot (y_2 - \bar{y})^2\}|$$

$$\underbrace{\leq}_{(*)} E\{|l_1| \cdot |y_1 - \bar{y}|^2 \cdot |l_2| \cdot |y_2 - \bar{y}|^2\} \underbrace{\leq}_{Lemma\ 1} \mathrm{E}\{|\bar{y}|^2 \cdot |\bar{y}|^2\} =$$

$$\mathrm{E}\{|\bar{y}|^4\} \underbrace{\to 0}_{Corollary} \quad \text{for } n \to \infty.$$

Inequality (*) rests upon the well-known relation $|E\{X\}| \leq E\{|X|\}$.
Since $|F_{3,n}| \to 0$ for $n \to \infty$, $F_{3,n}$ converges to 0 for $n \to \infty$.

Term $F_{1,n}$

$$\begin{aligned} &F_{1,n} \\ &= E\{m_1 \cdot m_2 \cdot y_1^2 \cdot y_2^2\} - E\{m_1 \cdot m_2 \cdot [y_1^2 \cdot 2 \cdot y_2 \cdot \bar{y} + y_2^2 \cdot 2 \cdot y_1 \cdot \bar{y}]\} \\ &\quad + E\{m_1 \cdot m_2 \cdot [y_1^2 \cdot \bar{y}^2 + y_2^2 \cdot \bar{y}^2]\} + 4 \cdot E\{m_1 \cdot m_2 \cdot y_1 \cdot y_2 \cdot \bar{y}^2\} \\ &\quad\quad -2 \cdot E\{m_1 \cdot m_2 \cdot [y_1 \cdot \bar{y}^3 + y_2 \cdot \bar{y}^3]\} + E\{m_1 \cdot m_2 \cdot \bar{y}^4\}. \end{aligned}$$

Since $y_1, \dots, y_n$ are independent and identically distributed and $E\{m_1 \cdot y_1^2\} \underbrace{=}_{(5)} \sigma_-^2$, it is obtained

$$\begin{aligned} F_{1,n} &= \sigma_-^4 - 2 \cdot 2 \cdot \underbrace{E\{m_1 \cdot m_2 \cdot y_1^2 \cdot y_2 \cdot \bar{y}\}}_{=:J_{1,n}} \\ &+2 \cdot \underbrace{E\{m_1 \cdot m_2 \cdot y_1^2 \cdot \bar{y}^2\}}_{=:J_{2,n}} + 4 \cdot \underbrace{E\{m_1 \cdot m_2 \cdot y_1 \cdot y_2 \cdot \bar{y}^2\}}_{=:J_{3,n}} \\ &-2 \cdot 2 \cdot \underbrace{E\{m_1 \cdot m_2 \cdot y_1 \cdot \bar{y}^3\}}_{=:J_{4,n}} + \underbrace{E\{m_1 \cdot m_2 \cdot \bar{y}^4\}}_{=:J_{5,n}}. \end{aligned}$$

It suffices to show that the terms $J_{1,n}, \dots, J_{5,n}$ converge to 0, to prove $F_{1,n} \to \sigma_-^4$.

- $J_{1,n}$:

$$0 \leq |J_{1,n}| = |E\{m_1 \cdot m_2 \cdot y_1^2 \cdot y_2 \cdot \bar{y}\}| =$$

$$\left| E\left\{\frac{1}{n} \cdot m_1 \cdot m_2 \cdot y_1^2 \cdot y_2 \cdot \left(y_1 + y_2 + \sum_{i=3}^{n} y_i\right)\right\}\right| =$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

$$\frac{1}{n}\cdot\left|E\{m_1\cdot m_2\cdot y_1^3\cdot y_2\}+E\{m_1\cdot m_2\cdot y_1^2\cdot y_2^2\}+\sum_{i=3}^{n}E\{m_1\cdot m_2\cdot y_1^2\cdot y_2\cdot y_i\}\right|$$

$$\underbrace{=}_{(2)}\frac{1}{n}\cdot\left|E\{m_1\cdot m_2\cdot y_1^3\cdot y_2\}+E\{m_1\cdot m_2\cdot y_1^2\cdot y_2^2\}\right|$$

$$\underbrace{=}_{(2)}\frac{1}{n}\cdot\left|E\{m_1\cdot y_1^3\cdot m_2\cdot y_2\}+\underbrace{E\{m_1\cdot m_2\cdot y_1^2\cdot y_2^2\}}_{(5)}\right|=\frac{1}{n}\cdot|v_3\cdot v_1+\sigma_-^4|\to 0$$

for $n\to\infty$.

- $J_{2,n}$:

$$0\le |J_{2,n}|$$
$$=|E\{m_1\cdot m_2\cdot y_1^2\cdot\bar{y}^2\}|\underbrace{\leqq}_{\substack{Cauchy-Schwarz\\ inequality}}\sqrt{\underbrace{E\{m_1^2\cdot y_1^4\}}_{\substack{=E\{m_1\cdot y_1^4\}=G_{4,n}\\ bounded}}\cdot\underbrace{E\{m_2^2\cdot\bar{y}^4\}}_{=E\{m_2\cdot\bar{y}^4\}=G_{0,n}\to 0}}$$
$$\to 0$$

for $n\to\infty$.

- $J_{3,n}$:

$$0\le |J_{3,n}|$$
$$=|E\{m_1\cdot m_2\cdot y_1\cdot y_2\cdot\bar{y}^2\}|\underbrace{\leqq}_{\substack{Cauchy-Schwarz\\ inequality}}\sqrt{\underbrace{E\{m_1^2\cdot m_2^2\cdot y_1^2\cdot y_2^2\}}_{=E\{m_1\cdot m_2\cdot y_1^2\cdot y_2^2\}=\sigma_-^4}\cdot\underbrace{E\{\bar{y}^4\}}_{\substack{\to 0\\ Lemma\ 2}}}$$
$$\to 0$$

for $n\to\infty$.

- $J_{4,n}$:

$$0\le |J_{4,n}|$$
$$=|E\{m_1\cdot m_2\cdot y_1\cdot\bar{y}^3\}|\underbrace{\leqq}_{\substack{Cauchy-Schwarz\\ inequality}}\sqrt{\underbrace{E\{m_1^2\cdot y_1^2\cdot\bar{y}^2\}}_{=E\{m_1\cdot y_1^2\cdot\bar{y}^2\}=G_{2,n}\to 0}\cdot\underbrace{E\{m_2^2\cdot\bar{y}^4\}}_{=E\{m_2\cdot\bar{y}^4\}=G_{0,n}\to 0}}\to 0$$

for $n\to\infty$

- $J_{5,n}$:

$$0\le J_{5,n}=E\{m_1\cdot m_2\cdot\bar{y}^4\}\le\underbrace{E\{\bar{y}^4\}}_{\substack{\to 0\\ Lemma\ 2}}\to 0\text{ for }n\to\infty.$$

Term $n \cdot (n-1) \cdot {c_n}^2$

$$n \cdot (n-1) \cdot c_n^2 \to 1 \text{ for } n \to \infty.$$

Summary

It is obtained:

$$F_{1,n} \to \sigma_-^4, F_n \to \sigma_-^4 \text{ and } n \cdot (n-1) \cdot c_n^2 \to 1.$$

These three results in combination lead to $E\{{t_n}^2\} \to \sigma_-^4$ (in Equation (14)).

References

Ang A., Chen J., Xing Yuhang, 2006, *Downside risk*, The Review of Financial Studies, 19 (2006), pp. 1191-1239.

Barrett G.F., Donald S.G., 2003, *Consistent tests for stochastic dominance*, Econometrica, 71 (2003), pp. 71-104.

Bawa V.S., Lindenberg E.B., 1977, *Capital market equilibrium in a mean-lower partial moment framework*, Journal of Financial Economics, 5 (1977), pp. 189-200.

Brooks Ch., Persand G., 2003, *The effect of asymmetries on stock index return Value-at--Risk estimates*, The Journal of Risk Finance, 4 (2003), pp. 29-42.

Cotter J., Hanly J., 2006, *Reevaluating hedging performance*, The Journal of Futures Markets, 26 (2006), pp. 677-702.

Davidson R., Duclos J., 2000, *Statistical inference for stochastic dominance and for the measurement of poverty and inequality*, Econometrica, 68 (2000), pp. 1435-1464.

Demirer R., Lien D., 2003, *Downside risk for short and long hedgers*, International Review of Economics and Finance, 12 (2003), pp. 25-44.

Demirer R., Lien D., Shaffer D.R., 2005, *Comparisons of short and long hedge performance: The case of Taiwan*, Journal of Multinational Financial Management, 15 (2005), pp. 51-66.

Hoechner B., Reichling P., Schulze G., 2017, *Pitfalls of downside performance measures with arbitrary targets*, International Review of Finance, 17 (2017), pp. 597-610.

Hoepner A.G.F., Oikonomou I., Sautner Z., Starks L.T., Zhou Xiaoyan, 2018, *ESG Shareholder Engagement and Downside Risk*, Working Paper, https://ssrn.com/abstract=2874252, downloaded on June 26, 2018.

Jin, Hanqing, Markowitz H.M., Zhou Xun Yu, 2006, *A note on semivariance*, Mathematical Finance, 16 (2006), s. 53-61.

Josephy N.H., Aczel A.D., 1993, *A statistically optimal estimator of semivariance*, European Journal of Operational Research, 67 (1993), pp. 267-271

Kotz S., Johnson N.L., 1985, *Multinomial Coefficients*, [in:] *Encyclopedia of Statistical Sciences*, Vol. 5, John Wiley & Sons, New York.

Lee, Hsiu-Chuan, Chien, Cheng-Yi, 2010, *Hedging performance and stock market liquidity: Evidence from the Taiwan futures market*, Asia-Pacific Journal of Financial Studies, 39 (2010), pp. 396-415.

Lewis A.L., 1990, *Semivariance and the performance of portfolios with options*, Financial Analysts Journal, 1990, July-August, pp. 67-76.

Markowitz H.M., 1959, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 17(23)

OPTYMALNE ESTYMATORY SEMIWARIANCJI: KOREKTA DOWODU JOSEPHY'EGO-ACZELA

Streszczenie: Semiwariancja jest intuicyjną miarą ryzyka, ponieważ koncentruje się na wartościach osiąganych poniżej celu, a nie na całkowitej zmienności. Aby z powodzeniem stosować semiwariancję w praktyce, potrzebny jest estymator semiwariancji; Josephy i Aczel podają takie oszacowanie. Niestety wspomniani autorzy nie udowodnili poprawnie asymptotycznej spójności i średniej kwadratowej błędów estymatora. Ich dowód zawiera bowiem błąd. Prezentowany artykuł koryguje błąd obliczeniowy w oryginalnym dowodzie Josephy'ego i Aczela i w ten sposób dostarcza badaczom i praktykom narzędzie m.in. do wyboru efektywnego portfela oraz pomiaru ryzyka.

Słowa kluczowe: analiza ryzyka, semiwariancja, estymacja statystyczna.