Finance Research Letters 48 (2022) 103000

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Averaging financial ratios☆

José Dias Curto, Pedro Serrasqueiro *

ISCTE Business School (IBS), Department of Quantitative Methods, Complexo INDEG/ISCTE, Av. Prof. Aníbal Bettencourt 1600-189 Lisboa, Portugal

ARTICLE INFO

Keywords:
Central tendency
Geometric average
Bootstrapping

ABSTRACT

Ratios represent a special kind of relation between two magnitudes, and computing the average of ratios is fairly common among academics and Finance practitioners. How should price-to-earnings (P/E) ratios be aggregated (averaged) at the portfolio level to provide a unified number? The arithmetic mean is the natural alternative. However, in case of financial ratios, it is generally accepted that the much less familiar harmonic mean may be more valuable, because it solves the upward bias encountered when using arithmetic mean. However, and to the best of our knowledge, there is no statistical evidence to show the superiority of the harmonic mean when computing the average of ratios. In this paper, by bootstrapping P/E ratios and earnings yield of companies listed in eight common stock indices, we compare the traditional averages and it is shown that geometric, not the harmonic average, as it is commonly accepted, is more suitable to average the ratios.

1. Introduction

Ratios represent a special kind of relation between two magnitudes, and are commonly employed in finance (Brown, 2011; Simoens and Vennet, 2020; Husna and Satria, 2019; Sriram, 2020; Husain and Sunardi, 2020; Gill et al., 2010; Agrrawal et al., 2010; Musallam, 2018). Estimation of the Sharpe ratio is an example (see Zsolt and Botond, 2020 for an application to cryptocurrencies), while distributional properties of the ratio of independent random variables have also been the subject of several studies (Nadarajah and Kotz, 2005; Weele, 2020). For example, how do firms' price-to-earnings (P/E) ratios aggregate at portfolio level to provide a unified number? The arithmetic mean is the natural alternative. However, there are often more suitable options for financial applications (Kaplan et al., 2010; Berishvili and Berishvili, 2020; Trönqvist et al., 1985; Martin and Stanford, 2007).

Three different situations arise when dealing with ratios:

1. **All Equal: numerators or denominators are the same across ratios**. Consider the average number of spectators per football match in the last five seasons. To get the exact value (the realized value, that we represent by "TA", the true average), we divide the sum of all spectators by the number of matches. Alternatively, as the data is available per year, we may compute the ratio per season and compute the average of ratios. Which average should one use? If we admit a constant number of matches (the denominator), the arithmetic average of ratios is exactly TA. Consider only two seasons, where the number of spectators is 301,000 and 400,400 and the number of matches is 70 per season. The average number of spectators per match is: $TA = (301000 + 400400)/(70 + 70) = 5010$. The two ratios per season are $R_1 = 301000/70 = 4300$, $R_2 = 400400/70 = 5720$ and the arithmetic average of the two ratios is exactly TA: $(4300 + 5720)/2 = 5010$.

☆ **Funding**: This work was supported by Fundação para a Ciência e a Tecnologia, grant UIDB/00315/2020.

* Corresponding author.
E-mail addresses: dias.curto@iscte-iul.pt (J.D. Curto), pnsoa@iscte-iul.pt (P. Serrasqueiro).

https://doi.org/10.1016/j.frl.2022.103000
Received 19 March 2022; Received in revised form 14 May 2022; Accepted 17 May 2022
Available online 25 May 2022
1544-6123/

2. **Become Equal: ratios have unequal numerators or denominators, but a constant value may be assumed for either**. If one invests in two stocks with prices of 400 and 200, and earnings per share are 20 and 40, respectively. The true average of price per unit of earnings (the value-weighted average, VWA) is: $VWA = (400 + 200)/(20 + 40) = 10$. The two P/E ratios are $R_1 = 400/20 = 20$, $R_2 = 200/40 = 5$ and the arithmetic and harmonic averages are 12.5 and 8, respectively. Neither average is 10, but we can buy two shares of the second stock and the investment is now the same for both stocks: 400. The true average of a portfolio with $1+2$ shares is $(400+400)/(20+80) = 8$, the value of the harmonic average of R_1 and R_2. According to Agrrawal et al. (2010), this is the main reason to use the harmonic mean: it has a more logical investment assumption, since most portfolios are neither on an equal number of shares per position basis, nor on equalized earnings per holding.
3. **All Different: numerators and denominators are different and may not be assumed constant**. Firstly, if individual numerators and denominators are available, it is possible to compute the true mean by summing all numerators and the denominators and obtain the ratio. Secondly, if only observed ratio values are available (individual numerators and denominators are unknown), we must learn from the previous case (when all data is available) to decide appropriately. By bootstrapping several samples from real financial data (see Section 3), we conclude that the geometric average, not the harmonic (as it is commonly accepted), is more suitable to estimate the true average of ratios.

Several unifying generalizations of the concept of average have been proposed in the literature, of which not only the three pythagorean but also the power mean are particular cases. Two such approaches are the Chisini mean (Chisini, 1929) and the Kolmogorov or generalized f-mean (see Kolmogorov and Castelnuovo, 1930; de Carvalho, 2016). While these approaches provide a more holistic notion of the concept of mean, their direct applicability is not always feasible, namely when dealing with financial ratios if the numerators and denominators of the ratios are unknown. We therefore focus on the power mean, not as an alternative, but rather as an application of the aforementioned generalizations, as shown in Section 2.

Our goal is to clarify the financial community about differences of ratio averaging methods. In Section 2 we shed light on the Chisini mean approach, focusing on the generalized (or power) mean. We then perform a bootstrap simulation study in Section 3, to infer about the most appropriate averaging method for financial ratios, namely P/E and Earnings Yield of eight financial datasets, and extend the analysis to the occurrence of stock splits. Additionally, we perform an out-of-sample analysis to further corroborate our hypothesis. Our results suggest that the geometric mean outperforms the other typical alternatives, especially in the presence of outliers.

2. The Chisini mean approach

According to Chisini (1929)[1] the mean of n homogeneous values $z_1, z_2, \dots, z_n$ with respect to the invariance requirement f, is the number (if it exists) $\bar{z}$ such that:

$$f(\bar{z}, \bar{z}, \dots, \bar{z}) = f(z_1, z_2, \dots, z_n). \tag{1}$$

Let $r_1, r_2, \dots, r_n$ be the ratios to be averaged:

$$r_1 = \frac{x_1}{y_1},\ r_2 = \frac{x_2}{y_2}, \dots, r_n = \frac{x_n}{y_n}. \tag{2}$$

A natural requirement to be satisfied, as the ratio includes two changing variables (x, y), is that the total x $(\sum x_i)$ and the total y $(\sum y_i)$ remain unchanged when replacing the observed ratios by their mean $(\bar{r})$. With respect to x $(x_i = r_i y_i)$, Eqs. (1) and (2) yield:

$$\sum_{i=1}^{n} x_i = r_1 y_1 + r_2 y_2 + \dots + r_n y_n = \bar{r} y_1 + \bar{r} y_2 + \dots + \bar{r} y_n \iff \bar{r} = \frac{\sum y_i r_i}{\sum y_i}, \tag{3}$$

leading to the weighted arithmetic mean $(\bar{r})$ (see first row of Table 1 in Graziani and Veronese, 2012).

With respect to y $(y_i = \frac{x_i}{r_i})$, the same Eqs. (1) and (2) yield:

$$\sum_{i=1}^{n} y_i = \frac{x_1}{r_1} + \frac{x_2}{r_2} + \dots + \frac{x_n}{r_n} = \frac{x_1}{\bar{r}} + \frac{x_2}{\bar{r}} + \dots + \frac{x_n}{\bar{r}} \iff \bar{r} = \frac{\sum x_i}{\sum \frac{x_i}{r_i}}, \tag{4}$$

which is the weighted harmonic mean $(\bar{r})$ (see third row of Table 1 in Graziani and Veronese, 2012). We note that, for constant y, Eq. (3) reduces to the simple arithmetic mean and, for constant x, Eq. (4) becomes the simple harmonic mean.

To compute the mean of the ratios based on Eq. (1), it is a requirement that individual numerator and denominator values are known. From Eqs. (3) and (4):

$$\bar{r} = \frac{\sum x_i}{\sum y_i} = \frac{r_1 y_1 + r_2 y_2 + \dots + r_n y_n}{\frac{x_1}{r_1} + \frac{x_2}{r_2} + \dots + \frac{x_n}{r_n}} = \frac{\bar{r}^2 \sum y_i}{\sum x_i} \to \bar{r}^2 = \frac{\frac{\sum r_i y_i}{\sum \frac{x_i}{r_i}}}{\frac{\sum y_i}{\sum x_i}}, \tag{5}$$

[1] The Chisini mean approach has been included by referee suggestion.

The practical problem we intend to address is how to compute average ratios at, for example, portfolio level, when individual components are either unavailable or impractical to obtain. One solution is to focus on the generalized (or power) mean, a particular case of the Chisini mean. In general, if $y = \phi(z)$ is a continuous and strictly increasing function, the Chisini mean is obtained from (1) as:

$$\bar{z} = \phi^{-1}\left[f(z_1, z_2, \dots, z_n)\right]. \tag{6}$$

For instance, if $f(z_1, z_2, \dots, z_n) = \phi(z) = \sum_{i=1}^{n} w_i z_i^k$, where w_i are nonnegative constants not all equal to zero, then we obtain from (6):

$$\bar{z} = \phi^{-1}\left(\sum_{i=1}^{n} w_i z_i^k\right) = \left[\frac{\sum_{i=1}^{n} w_i z_i^k}{\sum_{i=1}^{n} w_i}\right]^{\frac{1}{k}}, \tag{7}$$

for $z_i > 0, (k \neq 0, k \in \mathbb{R})$, which is the weighted power mean, a particular case of the Chisini mean. If $w_i = 1$, then the simple generalized or power mean with exponent k is:

$$\bar{z}_{GA}^{k}\left(z_1, z_2, \dots, z_n\right) = \left(\frac{1}{n}\sum_{i=1}^{n} z_i^k\right)^{\frac{1}{k}} \tag{8}$$

and the three Pythagorean means are special cases of $\bar{X}_{GA}^{k}$. For $k = -1, 1, 2$, we obtain the harmonic, the arithmetic, and the quadratic mean respectively. When $k \to 0$ we obtain the geometric mean. For recent applications of the generalized mean see, for example, Gou et al. (2019), Lu et al. (2020), Priam (2020) and Kolahdouz et al. (2020). In the bootstrap simulation study (see Section 3) the constant k is estimated to conclude about the appropriateness of each averaging method.

3. Bootstrapping P/E ratio and earnings yield

Bootstrapping is employed to compare the appropriateness of the three averages to estimate the realized portfolio P/E. Eight bootstrap samples were selected based on common stock indices: SP500 (421), NASDAQ (90), Stoxx600 (467), NIKKEI (171), FTSE 100 (74), SSEC (1282), CAC40 (32) and DAX30 (23) with non-negative P/E ratio on November 27, 2020. Data was obtained from http://www.investing.com.

A descriptive analysis of sample P/E and EY suggests that the empirical distributions of P/E and EY ratios are always positive asymmetrical and leptokurtic, pointing to the presence of outlying observations. The estimates for k, the exponent of the generalized or power mean in Eq. (8) that minimizes the squared error defined below in Eq. (9), seem far from −1, the value of the harmonic average. As most of k estimates are close to zero, we foresee that the geometric mean is the most appropriate to average the ratios under analysis. The two financial ratios are non-negative by definition, therefore the bootstrapping simulation study is conducted by splitting the ratios into two different categories: ratios above 1 (P/E) and ratios ranging from 0 to 1 (EY).

The value of the ratio results from the division of x by y: $r = \frac{x}{y}$. In the bootstrap simulation process we draw random samples from the eight original datasets with replacement 10,000 times. For each bootstrapped sample we get n values (sample size) for r (P/E and EY ratios), x (Price in P/E and Earnings per share in EY) and y (Earnings per share in P/E and Price in EY).

A useful general principle to consider when deciding which mean is more appropriate is to replace each observation by the mean and see which one produces the correct result in the context of the question being asked (see Eq. (1)). In this paper "the correct result" is the realized ratio: the division between the sum of individual numerators and denominators. By searching for the value of k that minimizes the Squared Error (SE)

$$SE = \left[\left(\frac{1}{n}\sum_{i}^{n} r_i^k\right)^{\frac{1}{k}} - rr\right]^2, \tag{9}$$

which is the squared difference between the average resulting from Eq. (8) and the realized ratio we are looking for the measure that is closer to the "correct result", the portfolio P/E realized value.

The objective is to find, for each sample, the average value that is closer to the realized value of each ratio. We end up with 10,000 estimates for k and obtain a 95% confidence interval based on the standard error of the sequence of simulated estimates. The routines have been programmed in R. Based on the limits of the central 95% bootstrapped confidence interval we draw conclusions about the best measure to average ratios. The results are summarized in Fig. 1.

To test whether the value of k is significantly different from $k = -1$ (harmonic), $k = 0$ (geometric) and $k = 1$ (arithmetic), firstly we calculate bootstrapped 95% confidence intervals for k. This enables us to distinguish between an estimate for k with a confidence interval including the values −1, 0 and 1. The geometric average is the most suitable if the CI contains 0 and excludes −1 and 1; similar reasoning applies to the remaining two averages: e.g. the harmonic is best if the CI contains −1 but and excludes 0 and 1.

For both P/E and EY, we do not reject the null that $k = -1$ and $k = 1$ in one of eight cases, respectively (DAX and CAC40, see Fig. 1). On the other hand, all confidence intervals contain $k = 0$ (the asymptotic k-exponent for the geometric mean): our simulation points to the geometric mean as the best method to average P/E and EY ratios, to the detriment of both the arithmetic and the harmonic.

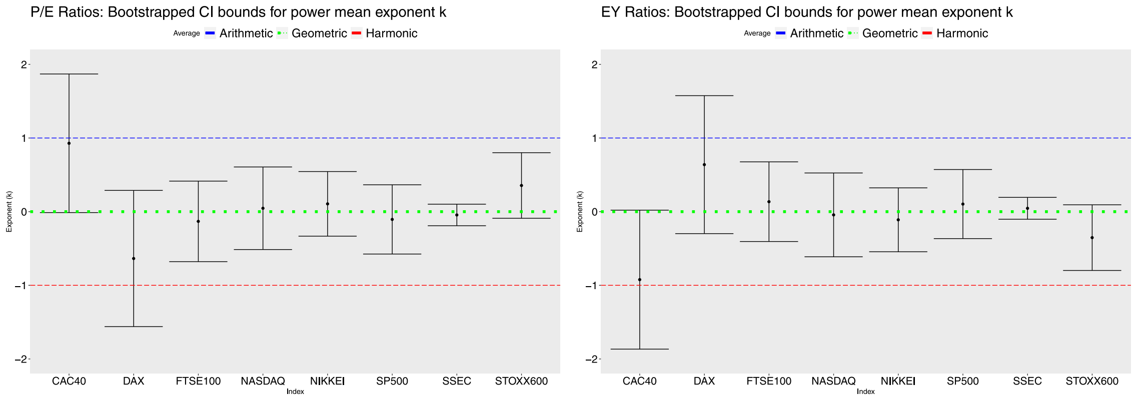

Fig. 1. Confidence Intervals for k and different stock indices. The charts contain the bootstrapped confidence intervals for power mean exponent k (see Eq. (8)) and each stock index. Horizontal lines depict the value of k that corresponds to the arithmetic (1, blue), geometric (0, green) and harmonic (−1, red). All confidence intervals contain $k = 0$, suggesting that the geometric is the most appropriate of the three pythagorean means to average P/E (> 1) and EY (< 1) ratios when individual components are unknown.

Table 1
Harvey–Newbold forecast outperformance test.

Comparison	HN test	p-value
Geometric vs. Arithmetic	−1.9905	0.9627
Geometric vs. Harmonic	−1.2513	0.8803
Harmonic vs. Arithmetic	−1.9247	0.9584

4. Additional diagnostics

We have performed an additional set of diagnostics to further validate our results.[2] To this effect, and to perform an out-of-sample analysis, we consider 483 SP500 companies with non-negative P/E ratios as of May 5th, 2022. Detailed results and simulation routines are available upon request to the authors.

Firstly, a $t-$test was conducted by bootstrapping 50 samples, to calculate the differences between each particular average (harmonic, arithmetic and geometric) and the realized P/E ratio of each sample. We state $H_0 : \mu_d = 0$ for each series of 50 differences, and obtain the corresponding p-values. Here, non-rejection of H_0 implies that estimates are significantly closer to the true average P/E ratio. Our results are confirmed, as we do not reject H_0 at $\alpha = 0.01$ for the geometric mean ($p = 0.017$), and reject H_0 for the harmonic and the arithmetic ($p = 0.000$).

Secondly, from a predictive viewpoint, that is, to the extent that an average is a function of past observations that can be used to predict future observations, an out-of-sample analysis was also conducted by applying the modified Diebold–Mariano test proposed by Harvey and Newbold (2000) (hereinafter HN test) to gauge whether each mean outperforms the other two (Curto and Pinto, 2012). We state that each particular mean encompasses its competitors as the null hypothesis. To compute the test we consider P/E ratios (as of May 5th, 2022) of eleven S&P 500 sectors (see also Curto and Serrasqueiro, 2021). The results of the HN test once again favor the geometric mean. We fail to reject the null that the geometric mean predictions encompass, or cannot be improved by combination with, the corresponding arithmetic and harmonic means at any reasonable significance level. Additionally, we conclude that the arithmetic mean seems to be the least appropriate (see Table 1).

The results also confirm the upward bias of the arithmetic average (where the lower bound of the CI exceeds the realized value) and the downward bias of the harmonic average (where the upper bound of the CI is lower than the realized value). Due to the presence of lower and upper outlying observations, neither the harmonic nor the arithmetic average seem to be the most suitable measures to average the P/E ratios. The simulation results suggest that, as the geometric average provides intermediate values and is not as influenced by outliers as the remaining two alternatives, it is more appropriate to estimate the true average of price per unit of earnings.

Thirdly, to conclude if more general functions can be optimized with respect to the type of averaging being employed, we conducted a simulation study to test if the arithmetic and geometric averages work best for arithmetic and geometric processes, respectively. Numerators and denominators of 1.000 ratios were generated from arithmetic and geometric Brownian motions (see Brătian et al., 2022; Ritschel et al., 2021; Stojkoski et al., 2020). We followed the same procedure as in Section 3, where we search for the value of k that minimizes (9). Here, the results are mixed and do not favor a particular average, depending on the starting value, the mean and the standard deviation of the simulated process. The confidence intervals include $k = 0$, corresponding

[2] The analyses in this section have been performed by referees' suggestions.

to the geometric average, but often include $k = -1$ and/or $k = 1$, regardless of the data generating process. More investigation is needed to shed further light on this topic.

5. Conclusions

Which averaging method should we use when computing the mean of the individual ratios? In finance, several authors argue that the harmonic mean is preferable, as the arithmetic mean assigns greater weight to higher and lower weight to lower ratios in samples. However, the harmonic mean tends strongly toward the smallest elements, as it tends (compared to the arithmetic mean) to mitigate the impact of large outliers. Hence, the arithmetic mean tends to overestimate and the harmonic mean tends do underestimate the true value of a mean ratio. On the other hand the geometric mean is the one that provides a better fit to the portfolio P/E realized value. We confirm this by presenting and discussing real data examples: the stock prices of companies listed in eight of the most influential stock indices. There are two special situations where one of the two means beats the other (and also the geometric average): if equal weight is assigned to the numerator's variable in each ratio, the harmonic mean is more suitable; when the same weight is given to the variable in the denominator, the arithmetic average provides the better result. If neither situation occurs, the geometric average, providing intermediate values, seems the most appropriate to estimate the true average of P/E and EY. Notwithstanding, as the "All Different" situation is more frequent than the remaining two (see Introduction), the geometric seems in fact most appropriate to average ratios, especially in the presence of lower and upper outlying observations. Further research can be done by extending comparison of averaging methods to financial time series data, and performing simulations of random processes (e.g. Geometric Brownian motion) with different distributions for innovations and dynamic functions to model drift parameters.

CRediT authorship contribution statement

José Dias Curto: Read and approved the manuscript. **Pedro Serrasqueiro:** Read and approved the manuscript.

References

Agrrawal, P., Borgman, R.H., Clark, J.M., Strong, R., 2010. Using the price-to-earnings harmonic mean to improve firm valuation estimates. J. Financial Education 36 (Fall/Winter).
Berishvili, V., Berishvili, L., 2020. Industry average financial ratios for Georgia. Ecoforum 9 (3(23)).
Brown, S., 2011. Averaging ratios: Characteristics of the error of approximation. World Appl. Program. 1 (5), 288–293.
Brătian, V., Acu, A.-M., Mihaiu, D.M., Şerban, R.-A, 2022. Geometric Brownian motion (GBM) of stock indexes and financial market uncertainty in the context of non-crisis and financial crisis scenarios. Mathematics 2022 (10), 309. http://dx.doi.org/10.3390/math10030309.
de Carvalho, M., 2016. Mean, what do you mean? Amer. Statist. http://dx.doi.org/10.1080/00031305.2016.1148632.
Chisini, O., 1929. Sul Concetto di media. Period. Mat. 2, 106–116.
Curto, J.D., Pinto, J.C., 2012. Predicting the financial crisis volatility. Econ. Comput. Econ. Cybernet. Stud. Res. J. 46 (1), 183–195.
Curto, J.D., Serrasqueiro, P., 2021. COVID-19 on S & P 500 sector indices and FATANG stocks volatility: An expanded APARCH model. Finance Res. Lett. http://dx.doi.org/10.1016/j.frl.2021.102247.
Gill, A., Biger, N., Tibrewala, R., 2010. Determinants of dividend payout ratios: Evidence from United States. Open Bus. J. 3, 8–14.
Gou, J., Ma, H., Ou, W., Zeng, S., Rao, Y., Yang, H., 2019. A generalized mean distance-based k-nearest neighbor classifier. Expert Syst. Appl. 115, 356–372.
Graziani, R., Veronese, P., 2012. How to compute a mean? The Chisini approach and its applications. Amer. Statist. 63 (1), 33–36.
Harvey, D., Newbold, P., 2000. Tests for multiple forecast encompassing. J. Appl. Econometrics 15, 471–482.
Husain, Sunardi, 2020. Firm's value prediction based on profitability ratios and dividend policy. Finance Econ. Rev. 2 (2).
Husna, A., Satria, Ibnu, 2019. Effects of return on asset, debt to asset ratio, current ratio, firm size, and dividend payout ratio on firm value. Int. J. Econ. Financ. Issues 9 (5), 50–54.
Kaplan, J., Fischer, D.G., Rogness, N.T., 2010. Lexical ambiguity in statistics: How students use and define the words: Association, average, confidence, random and spread. J. Stat. Educ. 18 (2).
Kolahdouz, F., Radmehr, F., Alamolhodaei, H., 2020. Exploring students' proof comprehension of the Cauchy generalized mean value theorem. Teach. Math. Appl. 39 (3), 213–235.
Kolmogorov, A.N., Castelnuovo, G., 1930. Sur la Notion de la Moyenne. G. Bardi, tip. della R. Accad. dei Lincei.
Lu, X.-M., Ma, Z., Zhang, C., 2020. Generalized-mean Cramér-Rao bounds for multiparameter quantum metrology. Phys. Rev. A 101, 022303.
Martin, W., Stanford, R.E., 2007. A methodology for estimating the maximum profitable turns for an ABC inventory classification system. IMA J. Manag. Math. 18 (3), 223–233.
Musallam, S., 2018. Exploring the relationship between financial ratios and market stock returns. Eurasian J. Bus. Econ. 11 (21), 101–116.
Nadarajah, S., Kotz, S., 2005. On the product and ratio of gamma and beta random variables. Allg. Stat. Arch. 89, 435–449.
Priam, R., 2020. Visualization of generalized mean estimators using auxiliary information in survey sampling. Comm. Statist. Theory Methods 49 (18), 4468–4489.
Ritschel, S., Cherstvy, A., Metzler, R., 2021. Universality of delay-time averages for financial time series: Analytical results, computer simulations, and analysis of historical stock-market prices. J. Phys. Complex. 2 (4), 045003.
Simoens, M., Vennet, R.V., 2020. Bank performance in Europe and the US: A divergence in market-to-book ratios. Finance Res. Lett. 101672.
Sriram, M., 2020. Do firm specific characteristics and industry classification corroborate voluntary disclosure of financial ratios: An empirical investigation of S & P CNX 500 companies. J. Manag. Gov. 24, 431–448.
Stojkoski, V., Sandev, T., Basnarkov, L., Kocarev, L., Metzler, R., 2020. Generalised geometric Brownian motion: Theory and applications to option pricing. Entropy 22, 1432. http://dx.doi.org/10.3390/e22121432.
Trönqvist, L., Vartia, P., Vartia, Y.O., 1985. How should relative changes be measured? Amer. Statist. 39 (1), 43–46.
Weele, T.J.V., 2020. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. Biometrics 76, 746–752.
Zsolt, N., Botond, B., 2020. Higher co-moments and adjusted sharpe ratios for cryptocurrencies. Finance Res. Lett. 101543.