

# COMPARING ENTROPY AND BETA AS MEASURES OF RISK IN ASSET PRICING

Galina Deeva<sup>1</sup>

<sup>1</sup>Department of Finance, Faculty of Economics and Administration, Masaryk University, Lipová 41a, 602 00 Brno, Czech Republic

## Abstract

DEEVA GALINA. 2017. Comparing Entropy and Beta as Measures of Risk in Asset Pricing. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 65(6): 1889–1894.

The paper establishes entropy as a measure of risk in asset pricing models by comparing its explanatory power with that of classic capital asset pricing model's beta to describe the diversity in expected risk premiums. Three different non-parametric estimation procedures are considered to evaluate financial entropy, namely kernel density estimated Shannon entropy, kernel density estimated Rényi entropy and maximum likelihood Miller-Madow estimated Shannon entropy. The comparison is provided based on the European stock market data, for which the basic risk-return trade-off is generally negative. Kernel density estimated Shannon entropy provides the most efficient results not dependent on the choice of the market benchmark and without imposing any prior model restrictions.

Keywords: entropy, risk measure, beta, asset pricing

## INTRODUCTION

In the recent review of emerging trends in asset pricing, Campbell (2015) considers entropy as a measure of uncertainty in the probability theory to be an alternative to variance in risk measuring. Even though the idea of using entropy in economic theory is at least 70 years old, it was famously ridiculed by Paul Samuelson and, given his authority, was popularized in econophysics just recently. Entropy was initially implemented in thermodynamics by Clausius (1870). Later Shannon (1948) showed that entropy concept can be applied in areas of science where probabilities can be determined. Since then, entropy became the major cornerstone of information theory (see Paninski 2006 for extensive overview), from which econophysics and modern financial economics borrow heavily.

In finance, entropy is viewed as “a measure of dispersion, a generalization of variance” (Backus *et al.* 2014). Maasoumi and Racine (2002) identify statistical properties of the entropy measure which are useful in regard to asset returns. Financial applications of entropy might be found in portfolio optimization (for example, Xu *et al.* 2011) or option pricing (Zhou *et al.* 2006). According to Backus *et al.*

(2014) entropy being a logarithmic measure can be easily computed for most of the popular asset pricing models usually defined as log-linear functions.

According to Ormos and Zibriczky (2014), implementing entropy in the asset pricing models allows dismissing the restriction on returns' normality distribution. This limitation was inflicted on asset returns by the use of standard deviation as a measure of uncertainty in the classic capital asset pricing models. Thus, restrictive assumptions of the CAPM are not applicable for returns in the model based on entropy risk measure. Moreover, calculation of entropy does not require defining any market portfolio.

In our study, we consider two risk measures which can be applied in asset pricing: the entropy – econophysics measure of risk, and the beta coefficient – covariance-variance ratio between the market portfolio and individual stock return. Beta coefficient has been the classic measure of risk in equilibrium based asset pricing models and had no powerful alternative to compete with. In accordance with results of previous research in this field we consider that entropy could be such an alternative. Unlike the variance that

measures concentration only around the mean (the mathematical introduction of beta is given in Equation 14), the entropy measures diffuseness of the density irrespective of the location of such concentration. In the statistical sense, the entropy is not a frequentist mean-centered measure, but the measure taking into account the entire empirical distribution without concentrating on a specific statistical moment. Simply put, the entropy is the weighted sum of all expected returns (next section provides a formal mathematical notation).

We are aimed to compare the explanatory power of standard CAPM beta and non-parametric ways of entropy calculations as risk measures arising from exposures to general market movements. Our study is closely related to Ormos and Zibriczky (2014), but distinct in two main considerations/contributions. First, instead of histogram-based density function estimation, we use kernel based and maximum likelihood estimations of entropy. Second, the explanatory power of risk measures is tested on extended sample of European stocks. European stock market is an interesting case for the testing of asset pricing models, since contrary to finance theory the basic risk-return trade-off is generally negative for European stocks (Aslanidis *et al.* 2016).

## MATERIALS AND METHODS

Since the concept of entropy is implemented in many scientific areas, there are several measures of the entropy which are used depending on the characteristics of data. In compliance with efficient market hypothesis, the stock market is an equilibrium system, which is a requirement for the implementation of Shannon and Rényi entropies. Even though the efficient market hypothesis is largely criticized, we leave the testing of entropy in terms of adoptive market hypothesis for further research.

Following Ormos and Zibriczky (2014) we consider a discrete random variable  $X$ , which has possible outcomes denoted by  $o_1, o_2, o_1, \dots, o_n$ , and corresponding probabilities  $p_i = \Pr(X = o_i)$ ,  $p_i \geq 0$  and  $\sum_{i=1}^n p_i = 1$ . The generalized entropy function for the variable  $X$  is defined as:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left( \sum_{i=1}^n p_i^\alpha \right) \quad (1)$$

where  $\alpha$  is the order of entropy,  $\alpha \geq 0$  and  $\alpha \neq 1$ , which expresses the weight put on each outcome. In case of  $\alpha = 1$ , the generalized entropy  $H_\alpha(X)$  converges into Shannon (1948) entropy:

$$H_1(X) = - \sum_{i=1}^n p_i \log p_i \quad (2)$$

In case of  $\alpha = 2$ , the generalized entropy  $H_\alpha(X)$  converges into Rényi (1961) entropy:

$$H_2(X) = - \log \left( \sum_{i=1}^n p_i^2 \right) \quad (3)$$

If the considered random variable  $X$  is continuous with a probability density function  $f(x)$ , the generalized continuous entropy (also known as differential) is defined as:

$$H_\alpha(X) = \frac{1}{1-\alpha} \int f(x)^\alpha dx \quad (4)$$

The Shannon and Rényi entropies in the continuous case are defined as follows:

$$H_1(X) = - \int f(x) \ln f(x) dx \quad (5)$$

$$H_2(X) = - \ln \int f(x)^2 dx \quad (6)$$

In practice the underlying probability function is unknown. The density function is estimated non-parametrically without assuming any particular theoretical probability distribution. The kernel-based density is estimated by the following formula:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K \left( \frac{x_o - x_i}{h} \right) \quad (7)$$

Where  $\hat{f}(x)$  is the density estimation of the random variable  $X$ ,  $n$  is the number of observations,  $h$  is the smoothing parameter (bandwidth) and  $K(\cdot)$  is the kernel function integrated to unity ( $\int K(u) du = 1$ ). The most common Gaussian form of the kernel function is used:

$$2\pi^{-1/2} \exp(-u^2) \quad (8)$$

The quality of the kernel function is largely based on the value of bandwidth. When a Gaussian kernel is used as a reference function, the optimal choice of bandwidth  $h$  or Silverman's rule of thumb is:

$$h = \left( \frac{4\hat{\sigma}^5}{3n} \right)^{1/5} \approx 1.06\hat{\sigma}n^{-1/5} \quad (9)$$

where  $\hat{\sigma}$  is the standard deviation of the sample.

As an alternative to kernel density estimation we also propose to use the maximum-likelihood estimator of Shannon entropy given by:

$$\hat{H}^{ML} = - \sum_{i=1}^n \hat{\theta}_i^{ML} \log(\hat{\theta}_i^{ML}) \quad (10)$$

The multinomial distribution is then used to connect the observed outcomes  $o_i$  with corresponding frequencies  $\theta_i$ :

$$\text{Prob}(o_1, o_2, \dots, o_n) = \frac{n!}{\prod_{i=1}^n o_i!} \prod_{i=1}^n \theta_i^{o_i} \quad (11)$$

The maximum likelihood estimator of  $\theta_i$  maximizes the function (11) for fixed number of outcomes  $o_n$  leading to the observed frequencies  $\hat{\theta}_i^{ML} = \frac{y_n}{n}$  with variances  $\text{Var}(\hat{\theta}_i^{ML}) = \frac{1}{n} \theta_i (1 - \theta_i)$

and zero bias  $\hat{\theta}_i^{ML} = 0$  as  $E(\hat{\theta}_i^{ML}) = \theta_i$ . Even though  $\hat{\theta}_i^{ML}$  is unbiased, the plug in entropy estimator  $\hat{H}^{ML}$  is not. According to Miller (1955), first-order bias correction leads to so called Miller-Madow estimator:

$$\hat{H}^{MM} = \hat{H}^{ML} + \frac{m > 0 - 1}{2n} \quad (12)$$

where  $m$  is the number of cells with  $o_i > 0$ .

The entropy-based risk measure is defined as follows:

$$\rho_i = \log H_\alpha(R_i - R_f) \quad (13)$$

where  $R_i$  is a stock return and  $R_f$  is a risk-free return. Given  $R_m$  as a market return, the CAPM beta risk measure is defined as follows:

$$\rho_i = \beta_i = \frac{\text{cov}(R_i - R_f, R_m - R_f)}{\sigma^2(R_m - R_f)} \quad (14)$$

As a simple way to compare the explanatory power of risk measures, we consider the linear relationship between the expected risk premium  $E(R_i - R_f)$  and risk measure  $\rho_i$ :

$$E(R_i - R_f) = a + b\rho_i + \varepsilon \quad (15)$$

Given the ordinary least square estimation of (15), the in-sample explanatory power of risk measure is denoted by the goodness-of-fit:

$$R^2(E(R_i - R_f), \rho_i) = \frac{\sum_{i=1}^n \left( E(R_i - R_f)_i - (\hat{a} - \hat{b}\rho_i) \right)^2}{\sum_{i=1}^n \left( E(R_i - R_f)_i - E(R_i - R_f) \right)^2} \quad (16)$$

### Data

The empirical analysis is based on European stock markets, for which STOXX® Europe 600 index is taken as a market benchmark. The index represents large, mid and small capitalization companies across 17 countries: Austria, Belgium, Czech

Republic, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom. The consistent prolonged data exists for 390 European stocks included in the index. The risk-free rate is the 10-year Germany government bond yield, reflected by Bloomberg in the automatic calculation of beta on the Bloomberg Terminal.

The daily data are obtained from Bloomberg for the period from the beginning of January 2003 to the end of June 2016. Each time series contains 3509 price observations. The dataset closely represents the statistical population of the most traded European stocks for which prolonged data exists. We evade from dividing the sample in country sub-samples for possible inconsistency of estimations based on smaller number of observations. Nevertheless, to adhere to normality restrictions of capital asset pricing model, we also consider weekly and monthly observations of the same data series.

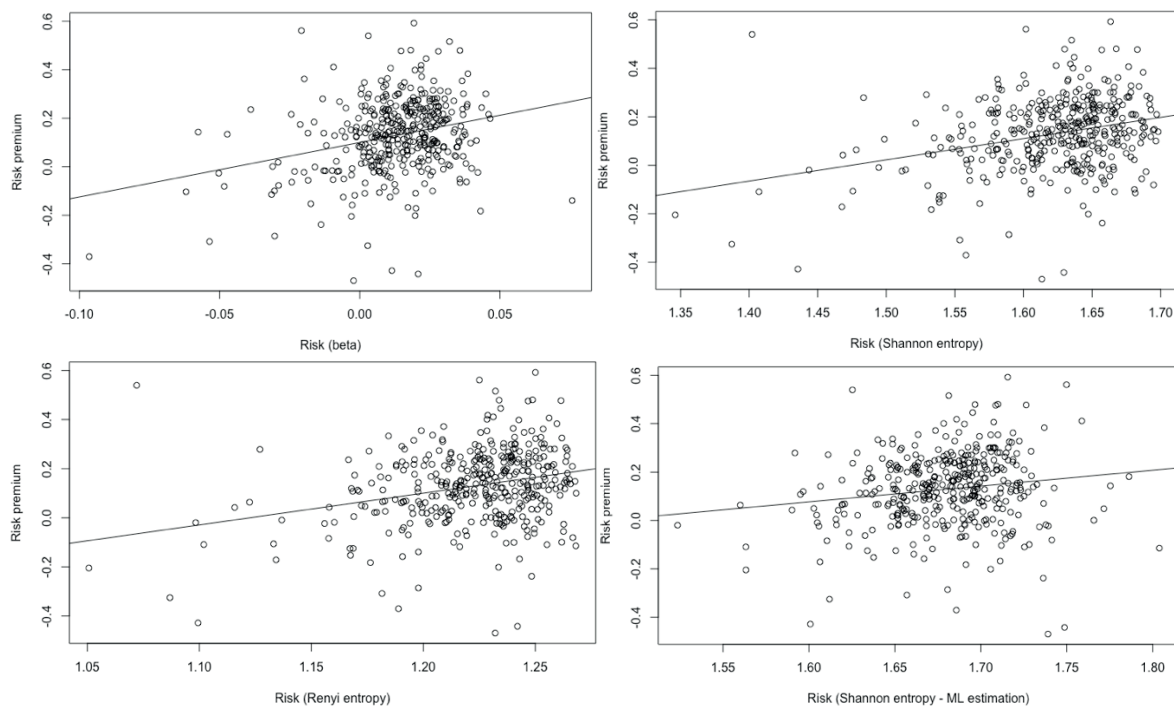
Tab. I reports the results of chosen normality tests (namely Shapiro-Wilk test, Anderson-Darling test and Pearson's chi-squared test). The results of normality tests for both market returns and risk-free returns indicate that examined time series do not follow normal distribution in any tested frequency, thus, violating the assumption behind capital asset pricing model and beta risk measure and justifying the use of non-parametric techniques.

### RESULTS

In order to analyse how the expected risk premium might be efficiently explained by risk measures, we estimate the risk for each security using the CAPM beta and three non-parametric entropy measures: kernel density estimated Shannon entropy, kernel density estimated Rényi entropy and maximum likelihood Miller-Madow estimated Shannon entropy. Fig. 1 shows the efficiency of explaining the average long-run risk premium by the considered risk measures (estimated on weekly data). There is certainly no strong linear relationship between risk premium and risk measure, even

I: Results of normality tests

Data	Test	Market returns		Risk-free returns	
		statistic	p-value	statistic	p-value
Daily	Shapiro-Wilk	0.9276	<2.2*10 <sup>-16</sup>	0.59845	<2.2*10 <sup>-16</sup>
	Anderson-Darling	51.62	<2.2*10 <sup>-16</sup>	364.44	<2.2*10 <sup>-16</sup>
	Pearson $\chi^2$	519.37	<2.2*10 <sup>-16</sup>	4040.6	<2.2*10 <sup>-16</sup>
Weekly	Shapiro-Wilk	0.91074	<2.2*10 <sup>-16</sup>	0.64156	<2.2*10 <sup>-16</sup>
	Anderson-Darling	7.9896	<2.2*10 <sup>-16</sup>	64.242	<2.2*10 <sup>-16</sup>
	Pearson $\chi^2$	79.685	1.278*10 <sup>-7</sup>	641.92	<2.2*10 <sup>-16</sup>
Monthly	Shapiro-Wilk	0.95451	4.492*10 <sup>-5</sup>	0.7477	2.783*10 <sup>-15</sup>
	Anderson-Darling	2.3364	6.108*10 <sup>-6</sup>	10.871	<2.2*10 <sup>-16</sup>
	Pearson $\chi^2$	29.8	0.005032	128	<2.2*10 <sup>-16</sup>



1: Explanatory Power of Risk Measures

though the clustering of the results is apparent. The observations for the trade-off between expected risk premium and risk measure are less dispersed along the regression line for kernel density estimated entropies.

Tab. II summarizes the results of linear regressions of risk premium on risk measure in three data frequencies: daily, weekly and monthly. The variability of observation frequency allows for the results to be robust to the level of relative volatility of the expected returns on individual stocks. For the daily data, the explanatory power of kernel density estimated Shannon entropy of 6,3% is the highest. Two other entropy measures perform similarly worse than Shannon entropy (goodness-of-fit of 5,1% and about 5,2%), but

better than the CAPM beta with 3,8% efficiency. The unexplained risk premium (also known as Jensen's alpha) given by the intercept in the regression is measured differently in sign by beta and entropies. Entropy-based alpha indicate no excessive returns given the amount of risk.

The risk measures are certainly sensitive to the choice of observations frequency. Shannon and Rényi entropies also perform better than CAPM beta in weekly-based and monthly-based regression settings. The performance of maximum likelihood Miller-Madow estimated Shannon entropy is, however, deteriorates with the diminishing observation frequency.

## CONCLUSION

Entropy is gaining prominence in asset pricing and financial modelling, but there are still uncertainties about the usage of the specific calculation techniques. Since there are several measures and estimation procedures of entropy, the question arises: which one should be used in asset pricing models. As a continuing discussion to previous studies, we considered three non-parametric possibilities. Our analysis based on European data demonstrates that the kernel density based entropy is a worthy option for the risk measure performing better than the CAPM beta. Even though we cannot directly compare our results to those of Ormos and Zibriczky (2014) set to different approaches in entropy calculations, our outcome supports the previous findings. The support comes in the form of the better explanatory power of Shannon entropy, which is the most widely used technique in information theory, as well as Rényi entropy. The performance of entropy-based measures is stable over the choice of observation frequency. Moreover, the entropy provides a risk measure not dependent on the choice of the market benchmark and without imposing any prior model restrictions. Therefore, the entropy produces more exact and stable measure of risk and should be implemented and assessed by practitioners in their investment decisions.

## II: In-sample explanatory power of tested risk measures

Data		Estimate	Std. error	t-statistic	p-value	R-squared
Daily	(Intercept)	0.019614	0.002251	8.712	$<2.2*10^{-16}$	0.03812
	Beta	0.625979	0.159627	3.922	0.000104	
	(Intercept)	-0.14287	0.09087	-2.271	$1.95*10^{-5}$	0.0633
	Shannon entropy	0.10970	0.06538	4.180	$4.80*10^{-7}$	
	(Intercept)	-0.15650	0.1498	-1.868	0.000103	0.05132
	Rényi entropy	0.15558	0.1348	3.025	$6.24*10^{-6}$	
	(Intercept)	-0.18966	0.11749	-1.963	$6.06*10^{-5}$	0.05206
	ML entropy	0.13660	0.07718	3.439	$5.32*10^{-6}$	
Weekly	(Intercept)	0.099892	0.009165	10.899	$<2.2*10^{-16}$	0.07161
	Beta	2.243827	0.410168	5.471	$8.06*10^{-8}$	
	(Intercept)	-1.2912	0.2324	-5.556	$5.13*10^{-8}$	0.08788
	Shannon entropy	0.8758	0.1432	6.114	$2.37*10^{-8}$	
	(Intercept)	-1.4532	0.2809	-5.174	$3.68*10^{-7}$	0.07565
	Rényi entropy	1.2946	0.2297	5.635	$3.37*10^{-8}$	
	(Intercept)	-0.9703	0.3578	-2.712	0.00699	0.02376
	ML entropy	0.6544	0.2129	3.073	0.00227	
Monthly	(Intercept)	0.58595	0.03997	14.66	$<2.2*10^{-16}$	0.01661
	Beta	2.01235	0.78609	2.56	0.0108	
	(Intercept)	-4.8580	1.4453	-3.361	0.000853	0.03603
	Shannon entropy	3.2440	0.8519	3.808	0.000163	
	(Intercept)	-5.843	1.858	-3.146	0.001785	0.03049
	Rényi entropy	5.125	1.467	3.493	0.000532	
	(Intercept)	-0.9828	2.5235	-0.389	0.697	0.001071
	ML entropy	0.9214	1.4286	0.645	0.519	

The predictive power of entropy measure of risk should also be considered in further studies alongside with different concepts of entropy, such as Tsallis entropy, Kullback cross-entropy and fuzzy entropy, which might be used for the formulation of entropy-based risk measure in terms of adaptive market hypothesis.

## Acknowledgements

The support of the Masaryk University internal grant MUNI/A/1039/2016 is gratefully acknowledged.

## REFERENCES

- ASLANIDIS, N., CHRISTIANSEN, C. and SAVVA, C. 2016. Risk-return trade-off for European stock markets. *International Review of Financial Analysis*, 46: 84–103.
- BACKUS, D., CHERNOV, M. and ZIN, S. 2014. Sources of entropy in representative agent models. *The Journal of Finance*, 69(1): 51–99.
- CAMPBELL, J. Y. 2015. Emerging Trends: Asset Pricing. In: SCOTT, R. and KOSSLYN, S. (Eds.) *Emerging Trends in the Social and Behavioral Sciences*. John Wiley & Sons, Inc.
- CLAUSIUS, R. 1870. On a mechanical theorem applicable to heat. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 40(1870): 122–127.
- MAASOUMI, E. and RACINE, J. 2002. Entropy and predictability of stock market returns. *Journal of Econometrics*, 107(1–2): 291–312.
- MILLER, G. A. 1955. Note on the Bias of Information Estimates. In: QUASTLER, H. (Ed.) *Information Theory in Psychology II-B*. Glencoe, IL: Free Press. pp. 95–100.
- ORMOS, M. and ZIBRICZKY, D. 2014. Entropy-Based Financial Asset Pricing. *PLoS ONE*, 9(12): e115742.
- PANINSKI, L. 2003. Estimation of entropy and mutual information. *Neural Computation*, 15(6): 1191–1253.
- RÉNYI, A. 1961. On measures of information and entropy. In: *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*. Vol. 1. Berkeley, Calif.: University of California Press, pp. 547–561.

- SHANNON, C. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27: 379– 423.
- XU, J. P., ZHOU, X. Y. and WU, D. D. 2011. Portfolio selection using  $\lambda$  mean and hybrid entropy. *Annals of Operations Research*, 185(1): 213–229.
- ZHOU, R. X., CHEN, L. M. and QIU, W.H. 2006. The entropy model of American bond option pricing. *Mathematics in Practice and Theory*, 36: 59–64.

Contact information

Galina Deeva: 385254@mail.muni.cz