

M Ű E G Y E T E M 1 7 8 2
Budapest University of Technology and Economics
PhD School in Business and Management

Dávid Zibriczky

Distribution-free non-parametric asset pricing

Theses

Supervisor: Mihály Ormos, PhD

BUDAPEST, 2016

Table of Contents

I	Introduction	1
II	Kernel-based Asset Pricing	7
	II.1 Single-factor models	7
	II.2 Multi-factor models	11
III	Entropy-based Asset Pricing	14
	III.1 Entropy as risk	14
	III.2 Empirical results	16
	References	19
	Related publications by the author	20

I Introduction

One of the most well known theories in finance is the Modern Portfolio Theory (MPT) that was developed in the 1950s (Markowitz, 1952). MPT attempts to capture the risk of an investment by the standard deviation of its return. Based on MPT's concept, the risk is characterized by two components: 1) a systematic or non-diversifiable risk that is inherent to the entire market or market segment; 2) idiosyncratic or non-systematic (specific) risk that relates to the company or a small group of similar companies, specifically. MPT states that diversification of a portfolio can reduce its risk, theoretically, all the idiosyncratic risk can be eliminated if large number of assets are involved to the portfolio; however, the presence of systematic risk cannot be avoided. The well-diversified portfolios maximize the expected return on a given risk level or minimize a risk to achieve an expected return. Based on MPT, these portfolios are called as efficient portfolios; they have systematic risk only and they are situated in a hyperbola (called Efficient Frontier) in the expected return – risk coordinate system. For a rational investor, MPT tells which efficient portfolio maximizes his utility based on his risk aversion. As an extension of MPT, if a risk-free asset is available, a combination of risk-free asset and a risky (tangency) portfolio results a line between risk-free and risky investment (called Capital Allocation Line). As MPT applies standard deviation to capture risk, it assumes normal distribution of asset returns; however, several studies show that this assumption does not hold for daily returns.

As an extension of Modern Portfolio Theory, Capital Asset Pricing Model (CAPM) was introduced in the 1960s (Treyner, 1962; Sharpe, 1964; Lintner 1965a,b; Mossin, 1966). In asset pricing models, we are about to find some kind of equilibrium between risk and return, no matter how the risk is defined. In the case of CAPM equilibrium, the model assumes the existence of an efficient market portfolio that includes all available risky assets in the market. As the market portfolio is efficient, it is also situated in Efficient Frontier. Based on CAPM, risk is characterized by the beta parameter that is the relative sensitivity of the volatility of asset (or portfolio) returns to the market portfolio. The model assumes linear relationship between asset returns and market returns that is characterized by Characteristic Line (CL); beta is the slope of this line. As CAPM assumes that the investors are rational and they hold efficient portfolios, beta captures the systematic risk of an asset only. CAPM states that the expected return is a linear function of beta exclusively and this relationship is characterized by Security Market Line (SML). Beta has a good interpretation, because it expresses the relative variance to the market; if the beta of an asset is greater than one, it is riskier than the market portfolio, and vice versa. CAPM assumes positive slope of SML that means an investor expect higher

returns taking higher risk. As it is simple to estimate a risk and expected return of an asset, CAPM is popular in financial analysis; however, it has received several criticisms.

In our research, we aim to avoid the assumption of 1) linearity between expected return and risk, 2) linearity between return of an asset and market return, 3) normal distribution of returns and 4) the existence of the market portfolio itself. The goal of this dissertation is to derive an application of distribution-free non-parametric model in risk estimation and asset pricing. As we have introduced, CAPM has several theoretical assumptions that may not hold in real life circumstances. Based on his empirical tests, Jensen (1968) introduces a performance index (Jensen's alpha), which explains abnormal returns over the risk adjusted (normal) return. Jensen's alpha is also the constant value of risk premium that cannot be explained by Characteristic Line, in other words, the constant coefficient of cross-sectional linear regression. For precise estimation of risk, it seems to be necessary to have linear relationship between the return of assets and the market return; otherwise, the standard linear estimator methods (e.g. Ordinary Least Squares) may calculate biased slope and intercept of linear regression. In this dissertation, we introduce a univariate non-parametric kernel regression method that is capable to characterize the relationship between risk and return and the relationship between the return of assets and market returns, even if the linearity assumption does not hold. Based on the goodness of fit of regression models, we show that kernel regression outperforms the linear one in all cases, thus it is also capable to estimate risk and abnormal performance. Using non-parametric regression, we deduce a hypothesis testing method to decide if the assumption of linearity is valid for CL. We show that the linearity can be rejected for U.S. stocks at 95% confidence level; therefore, we introduce an alternative non-linear estimation of risk and abnormal performance. We also show that asset returns can be explained by third-degree polynomial of market returns if the linearity does not hold. Comparing the linear- and kernel-based estimation of beta, we show that risk is significantly underestimated by CAPM if the linearity does not hold. We find that linearity is more likely rejected for risky assets.

As CAPM assumes linear relationship between the expected return and beta, we also investigate this assumption. We show that non-parametric beta is different only if the linearity does not hold, which confirms the consistency of non-parametric estimation with linear methods; therefore, we apply non-parametric betas for the estimation of Security Market Line. We show that the hypothesis of linearity for the Security Market Line cannot be rejected at any usual significance level. Based on the investigation of the slope and intercept of SML by market capitalization, we find that the slope of SML in the segment of small companies is negative.

The interpretation of this result is that lower returns is expected by taking higher risk that contradicts the theory of risk premium. We find that intercept of the SML is significant for small companies even if the risk is estimated by non-parametric beta; furthermore, we measure the highest expected risk premium for small companies, which confirms the small firm effect (Banz, 1981; Basu, 1983). According to these results, our first thesis is the following:

Thesis 1 (Erdős et al., 2010a,b; Erdős et al., 2011): The assumption of linearity for the Characteristic Line of CAPM can be rejected for U.S. stocks. The risk is significantly underestimated by CAPM beta for those stocks where linearity does not hold. On the other hand, the linearity of Security Market Line (SML) cannot be rejected; however, the slope of the SML for small companies is negative that contradicts the theory of risk premium in CAPM.

In order to improve the explanatory power and to answer the undermining anomalies of single-factor CAPM, several multi-factor models are introduced. One of the most well-known extension or reformulation of CAPM is the Fama-French three-factor model (Fama and French, 1996) that attempts to explain the returns of asset by 1) market returns, 2) the difference in returns between stocks of small and large market capitalization companies (SMB) and 3) the difference in returns between stock with high book-to-market ratios minus low ones (HML). Based on the work of Fama-French (1992 and 1996), Carhart (1997) extends their model with the momentum parameter (MOM) that is an empirically observed tendency for persistency meaning that rising asset prices to rise further, and falling prices to keep falling. His model is known as Carhart four-factor model. Both models hold the main assumptions of CAPM.

Similarly to the analysis we run on the CAPM, we also investigate the linear assumptions of the aforementioned models on explanatory factors. We deduce a multivariate non-parametric kernel regression to explain asset returns in a non-parametric way. As we don't know the real relationship between the asset returns and explanatory factors, we approximate this by multivariate kernel regression and discuss a multivariate hypothesis testing the linearity. We also deduce non-parametric estimation of coefficients to compare them with the standard estimations. Based on results of the hypothesis testing, we show that linearity cannot be rejected for neither Fama-French model nor Carhart model in any investigated segments at any usual confidence level. Although the linearity holds, we find that the coefficient of HML factor of the models is significantly overestimated by linear methods. Based on the factor analysis, we present that the coefficient of SMB factor is negatively, the coefficient of MOM factor is

positively correlated to the size of the company. Considering these results, we formulate the following thesis:

Thesis 2 (Erdős et al., 2011): The extension of CAPM by the Fama-French factors is capable to explain the returns of U.S. stocks by linear regression; therefore, linear estimation of Fama-French risk coefficients is adequate.

The second direction of our research is the investigation of entropy as an alternative non-parametric method that is capable to characterize risk by non-normal returns. Entropy is a mathematically defined quantity that is generally used for characterizing the probability of outcomes in a system that is undergoing a process. Originally, Clausius (1870) introduces entropy to characterize the disorder of heat in an isolated system in thermodynamics. In statistical mechanics, the interpretation of entropy is the measure of uncertainty about the system that remains after observing its macroscopic properties (pressure, temperature or volume). In information theory, entropy quantifies the expected value of the information in a message or, in other words, the amount of information that is missing before the message is received. The more unpredictable the message that is provided by a system, the greater is the expected value of the information contained in the message. As entropy characterizes the unpredictability of a random variable, our conjecture is that it can also be implemented to capture financial risk of an investment. Based on our approach, we apply continuous (differential) entropy on the returns of assets to characterize their risk. Higher entropy means higher uncertainty in returns that is interpreted as risk of an asset. The differential entropy of a random variable is similar to its standard deviation if it is normally distributed. Several studies show that daily return of assets follow non-normal distribution; therefore, standard deviation as a risk measure of MPT cannot capture the risk of an asset properly. An advantageous property of entropy is that it is distribution-free as it is estimated by distribution-free methods. We argue that eliminating the assumption of normality, entropy can capture the risk of an asset more accurately.

In our analyses, we discuss two types of entropy function – the Shannon- and Rényi entropy – and three types of estimation methods, the histogram-, the sample spacing- and the kernel-based estimation. We analytically show that entropy-based risk measure satisfies the axiom of positive homogeneity; furthermore, it satisfies the axiom of subadditivity and convexity if the distribution of the return of portfolios is normal. However, it does not satisfy the axiom of translation invariance and monotonicity; therefore, it is not a coherent risk measure (Artzner et al., 1999). Although the entropy-based risk measure is not coherent, we show that

it can be used for asset pricing efficiently. We propose an evaluation procedure that measures the linear explanatory and predictive power of risk measures in short- and long term. Based on our results, among the entropy estimation methods, histogram-based estimation offers the best tradeoff between explanatory and predictive power, therefore we deduce a simple formula of the estimation. Among entropy functions, we show that Shannon entropy has better short-term explanatory and predictive power, Rényi entropy is more accurate in long term. Based on long-term empirical results, we find insignificant intercept and significant coefficient of regression line of entropy-based risk measures, which concludes that entropy is capable to explain expected risk premium on its own. We also evaluate standard deviation of MPT and beta of CAPM as baseline methods. We find that Shannon entropy overperforms both standard risk measures and it is more reliable than CAPM beta; however, if the market trend becomes visible, we measure mixed results. Extending our methodology to multivariate risk analysis, we show that entropy can also be used as an extension of multi-factor asset pricing models, primarily for less-diversified portfolios. According our results, our third thesis is the following:

Thesis 3 (Ormos and Zibriczky, 2014): Entropy of the risk premium is an efficient measure of risk of assets on the capital markets. Entropy is more accurate on explaining (in-sample) and predicting (out-of-sample) returns than standard deviation or CAPM beta.

As mentioned above, differential entropy is similar to standard deviation if the distribution of returns is normal. More precisely, we can construct a formula for differential entropy that differs from standard deviation only in a constant. As the returns follows non-normal distribution we expect different risk measures applying the entropy; however, we measure similar behavior to standard deviation of MPT, more specifically, (1) capturing systematic and idiosyncratic risk (2) characterizing diversification effect and (3) efficient portfolios are situated on a hyperbola in the expected return – risk coordinate system. We empirically show that the entropy-based risk measure satisfies the axioms of subadditivity and convexity for any two portfolios with 99% confidence. We show that the average entropy of a random single-element portfolio decreases by 40% if 10 assets are involved instead of one. Based on that, we formulate our last thesis:

Thesis 4 (Ormos and Zibriczky, 2014): Entropy captures both systematic and idiosyncratic risk of an investment. It is capable to characterize the effect of diversification; the expected entropy of a portfolio decreases by the number of securities involved.

Formulating the dissertation, we discuss non-parametric models in two chapters. In Chapter II, we introduce kernel regression. First, we deal with univariate regression models and propose univariate hypothesis testing procedure to decide whether the assumptions of linearity of CAPM holds. We introduce a non-parametric estimation of abnormal return and risk. After the introduction of univariate methodology, we evaluate linear and non-parametric methods on S&P dataset. Second, we extend our methodology to test multi-factor models, namely Fama-French three-factor and Carhart four-factor models and we perform the same evaluation as for CAPM. With the introduction of multi-factor hypothesis testing, we present the results of univariate polynomial testing procedure as well.

In Chapter III, we introduce entropy as risk measure. First, we introduce discrete and differential entropy. We discuss the most frequently applied entropy functions and estimation methods. As we consider histogram-based entropy function the most efficient estimation method, we deduce a simple built-in formula for Shannon- and Rényi entropy function. We also investigate whether the entropy-based risk measure is coherent. After that, we propose an evaluation methodology to measure the explanatory- and predictive power of risk measures. We investigate how entropy behaves if we diversify the portfolios and we compare the results of entropy-based risk measures to standard deviation and CAPM beta as baseline measures. Finally, as an outlook, we evaluate Fama-French three-factor model, Carhart four-factor model, higher moments and their combination with entropy.

II Kernel-based Asset Pricing

Linear asset pricing is adequate only if the relationship between risk and return is linear. We discuss a hypothesis testing method to decide if the linearity holds between return (risk premium) and risk factors and between the daily return of assets and the daily return of market portfolio. For testing linearity, we introduce a non-parametric kernel regression as an approximation of real regression function. We test the hypothesis of linearity of CAPM (Characteristic Line and Security Market Line) with univariate setting of non-parametric method; furthermore, we test the linearity of multi-factor models (Fama-French three-factor model and Carhart four-factor model) by multivariate extension of introduced methodology. If the linearity does not hold, then the estimated parameters by linear methods are biased and inconsistent; therefore, we introduce an alternative non-parametric estimation of risk coefficients and abnormal performance. In our study, we pick daily returns of 50-50 randomly chosen stocks from various index of the S&P universe (from the S&P 500, S&P MidCap 400, S&P SmallCap 600) using data from the Center of Research in Stock Prices (CRSP) database for the period of 1999 to 2008. We apply daily logarithmic returns of stocks, risk free asset and the proxy of market portfolio for CAPM; furthermore, we use daily observations of SMB, HML and momentum factors for multi-factor models.

II.1 Single-factor models

Nadaraya (1964) and Watson (1964) introduce a univariate non-parametric regression between continuous random variable X and Y that applies kernel function – called Nadaraya-Watson estimator – in the following formula

$$\hat{y} = \hat{m}_h(x) = \sum_{i=1}^n W_{hi}(x) y_i, \quad (\text{II.1})$$

where the Nadaraya-Watson weighting function is

$$W_{hi}(x) = \frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}, \quad (\text{II.2})$$

and \hat{y} is an estimation in the function of x , y_i and x_i is an observation of Y and X , K is the kernel function and h is the bandwidth, that should be optimally selected. As the selection of kernel function K is only of secondary importance (Härdle et al., 2004), we use Gaussian kernel

function, because it is differentiable at every point and we expect smoother derivative function. In order to find the optimal value of bandwidth h , we optimize the following generalized cross-validation penalized objective function:

$$CV(h) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{m}_h(x_i))^2 \left(1 - \frac{1}{n} W_{hi}(x_i)\right)^{-2}. \quad (\text{II.3})$$

For the minimization of $CV(h)$, we use simplex-search method with a properly chosen initial value. Cross-validation ensures that we avoid overfitting the kernel regression model. In order to compare the accuracy of non-parametric regression with the linear model, we use R^2 as the measure of goodness of fit.

Kernel regression is introduced for providing an approximation of real regression function that is unknown. In order to test the hypothesis of linearity for Characteristic Lines and Security Market Line, we describe a hypothesis testing procedure for any parametric regression $m_\theta(x)$ based on the work of Härdle et al. (2004). The null hypothesis is $H_0 : m(x) = m_\theta(x)$, which is tested against the alternative hypothesis $H_1 : m(x) \neq m_\theta(x)$. We approximate $m(x)$ with kernel regression $\hat{m}_h(x)$. We use an error function between non-parametric and Nadaraya-Watson weighted parametric regression as

$$T = \sqrt{h} \sum_{i=1}^n \left(\hat{m}_h(x_i) - \sum_{j=1}^n W_{hi}(x_j) m_{\hat{\theta}}(x_j) \right)^2. \quad (\text{II.4})$$

As the distribution of T is unknown, we apply wild bootstrapping to generate random samples. In each iteration, we generate $y_i^* = m_\theta(x_i) + \varepsilon_i^*$ dependent variable for $i=1,2,\dots,n$, where ε_i^* is generated by the residual $\hat{\varepsilon}_i$ of original regression and the rule of ‘‘Golden ratio’’. We estimate parameters θ^* and calculate T^* , in the same way as T , applying k_b iterations. As we apply one-tail test for T^* , H_0 is rejected at significance level α , if $\Pr(T > T^*) \geq (1 - \alpha)$.

Figure II.1 shows the Characteristic Curve (estimated by kernel regression) with corresponding confidence bands and Characteristic Line (estimated by linear regression) for two example securities (Panel A: linear; Panel B: non-linear). Panel B suggests that linearity is primarily rejected in the tail of the distribution; however, we may also find statistically significant differences close to the core.

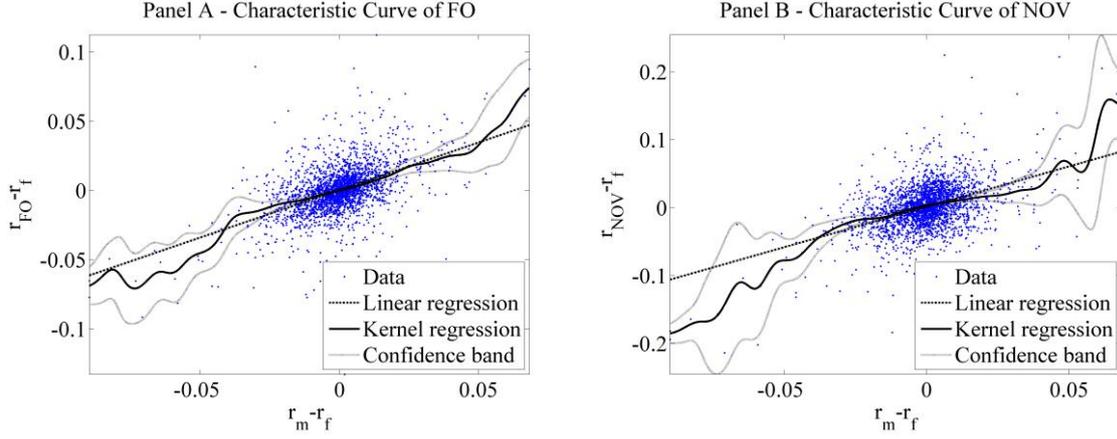


Figure II.1. The Characteristic Curve and Characteristic Line of example stocks

Notes: The figures show the Characteristic Curve (CC) and Characteristic Line (CL) of a company where the linearity of CL holds (“FO”, Panel A) and a company where the linearity of CL is rejected (“NOV”, Panel B). We estimate the CC and CL of both companies by two different methods: (1) Kernel regression (bold curve) between the risk premium of stock and risk premium of market portfolio, using Gaussian kernel and the Nadaraya and Watson weighting function. The grey curves represent the 95% confidence bands. (2) Linear regression (dashed line) between the risk premium of stock j and risk premium of market portfolio.

If the linearity of the Characteristic Line is rejected, the linear estimation of risk and abnormal returns may be biased; therefore, an alternative non-linear estimation of these should be applied.

We approximate the derivatives of $m(x)$ by kernel-weighted least squares

$$\hat{\boldsymbol{\beta}}(x) = \left(\hat{\beta}^{(0)}(x), \hat{\beta}^{(1)}(x), \dots, \hat{\beta}^{(p)}(x) \right)^T = \left(\mathbf{P}^T \mathbf{W} \mathbf{P} \right)^{-1} \mathbf{P}^T \mathbf{W} \mathbf{y}, \quad (\text{II.5})$$

where $\hat{\beta}^{(j)}(x)$ is the j^{th} derivative estimation of $m(x)$, \mathbf{P} is an $n \times (p+1)$ matrix generated from the observation of explanatory variable in order $j=0,1,\dots,p$, \mathbf{W} is a weight matrix and \mathbf{y} is the observation of dependent variable. The non-parametric beta (risk) is estimated by

$$\hat{\beta}_{KR} = \frac{1}{n} \sum_{i=1}^n \hat{\beta}^{(1)}(x_i) \quad (\text{II.6})$$

and the estimation of non-parametric alpha (abnormal return) is the following:

$$\hat{\alpha}_{KR} = \frac{1}{n} \sum_{i=1}^n y_i - \hat{\beta}_{KR} x_i \quad (\text{II.7})$$

We apply linearity testing of Characteristic Lines (CL) and non-parametric estimation of risk and abnormal performance for all securities. Table II.1 summarizes the average results aggregated by market capitalization (S&P 500, S&P MidCap 400 and S&P SmallCap 600) and for all companies. We see that kernel regression outperforms linear regression in all segments.

We show that linearity of CLs is rejected in segment S&P 500 and for all companies, because the linearity of CLs is rejected more than 5% of the included stocks.

Table II.1. Summary of linearity testing of Characteristic Lines, alpha- and beta estimation

Segment	N	$N(H_1)$	$P(H_1)$	$\overline{R_{KR}^2}$	$\overline{R_{LR}^2}$	$\overline{\hat{\alpha}_{KR}}$	$\overline{\hat{\alpha}_{LR}}$	$\overline{\hat{\beta}_{KR}}$	$\overline{\hat{\beta}_{LR}}$
S&P 500	50	10	20.0%	0.264	0.244	0.041	0.042	0.972	0.975
S&P MidCap 400	50	2	4.0%	0.243	0.224	0.056	0.055	0.962	0.956
S&P SmallCap 600	50	2	4.0%	0.187	0.171	0.070	0.069	1.010	0.924
All companies	150	14	9.3%	0.231	0.213	0.056	0.055	0.981	0.952
All companies H_0	136	0	0%	0.227	0.209	0.055	0.054	0.947	0.926
All companies H_1	14	14	100%	0.276	0.252	0.064	0.063	1.316	1.206

Notes: In the table, we show the average of estimated parameters and statistics of Characteristic Curve and Characteristic Line of 50-50 randomly chosen companies from the S&P 500, S&P MidCap 400 and the S&P SmallCap 600 universe, respectively. We estimate the Characteristic Curve of stock by kernel regression, using Gaussian kernel and the Nadaraya and Watson weighting function; and we estimate the Characteristic Line by linear regression between the series of risk premium of stock and risk premium of market portfolio. The 1st column is the segment in we aggregate the results. H_1 indicates the subgroup of all companies, where the linearity of Characteristic Line is rejected at confidence level 95%, H_0 indicates the rest of the companies. The 2nd and 3rd column shows the number of assets in the investigated segment (N) and the number of assets where the linearity is rejected. The 4th column is the ratio of the number of assets where the linearity is rejected and the all assets. The following columns shows the average goodness of fit, alpha and beta estimated by kernel- and linear regression, alternately.

Aggregating the results for companies where the linearity holds (H_0) and where it is rejected (H_1), we find the followings. Beta is significantly underestimated by linear regression if the linearity is rejected; otherwise, the difference between the linear and non-parametric beta is not significant. Furthermore, we also find that linearity is more likely rejected for risky assets. These statements are validated by paired two-sample t -tests, which is summarized in Table II.2.

Table II.2. Two-sample t -test of average beta in various subgroups

Group 1	N	Mean	Var.	Group 2	N	Mean	Var.	t	p	Sig.
Kernel beta in H_0	136	0.9468	0.1279	Linear beta in H_0	136	0.9257	0.0948	1.86	0.0629	
Kernel beta in H_1	14	1.3157	0.1874	Linear beta in H_1	14	1.2064	0.1173	2.39	0.0326	*
Kernel beta in H_1	14	1.3157	0.1874	Kernel beta in H_0	136	0.9468	0.1279	3.08	0.0076	**

Notes: The table summarizes the results of two-sample t -test between various subgroups. Column 1-4 and column 5-8 contains the (1) short name of the subgroups (2) the number of securities in the group (3-4) the average value and variance of beta in the subgroup. H_1 and H_0 indicates the group whether the linearity of the Characteristic Line is rejected or not, Linear beta and Kernel beta shows if the beta is estimated by linear and kernel regression. The last three columns are: the test statistic t , the p -value and the significance level of the test. ** and * indicates significance level 1% and 5%, respectively. If t is positive (or negative) and the test is significant, then the mean of Group 1 is significantly higher (or lower) than mean of Group 2.

For hypothesis testing of the linearity of Security Market Lines (SML), we estimate non-parametric beta for all securities. Table II.3 shows the result of linearity testing, the goodness

of fit and estimated parameters of the SMLs by market capitalization. We see that linearity cannot be rejected for any segments (p -values are higher than 0.05); however, kernel regression shows higher goodness of fit. Investigating the slope ($\hat{\beta}'$) of SMLs, we find that it is negative for small companies, which contradicts the theory of risk premium in CAPM; furthermore, we measure significant intercept ($\hat{\alpha}'$) of SML, which concludes that CAPM beta cannot explain expected risk premium on its own. Besides that, we also measure the highest expected risk premium for small companies, which confirms the small firm effect (Banz, 1981; Basu, 1983).

Table II.3. Estimation of Security Market Lines

Segment	N	$E(r - r_f)$	p	R_{KR}^2	R_{LR}^2	$\hat{\alpha}'_{KR}$	$\hat{\alpha}'_{LR}$	$\hat{\beta}'_{KR}$	$\hat{\beta}'_{LR}$
S&P 500	50	0.0357	0.144	0.117	0.052	0.028	0.007	0.008	0.029
S&P MidCap 400	50	0.0483	0.728	0.131	0.104	0.009	0.001	0.041	0.049*
S&P SmallCap 600	50	0.0548	0.760	0.089	0.026	0.087	0.081**	-0.034	-0.028
All companies	150	0.0462	0.688	0.042	0.010	0.037	0.031*	0.009	0.015

Notes: We estimate Security Market Lines (SMLs) for the S&P 500, S&P MidCap 400, S&P SmallCap 600 stocks and for all the stocks in our database. The rows 1,2,3,4 include the estimated parameters and statistics of the SMLs for S&P 500, S&P MidCap 400, S&P SmallCap 600 indices and for all companies, respectively. We estimate the SMLs with kernel regression and linear regression between the expected risk premium and the non-parametric beta of the stocks using Gaussian kernel and the Nadaraya and Watson weighting function for kernel regression. The 1st column represents the segment, the 2nd is the number of stocks involved in the segment. The 3rd column shows the mean return of the given segment for the estimation period, column 4 is the p -value of the linearity test. $\hat{\alpha}'_{KR}$ and $\hat{\alpha}'_{LR}$ are the estimation of constant value (abnormal returns) of Security Market Line (Curve) using kernel and linear regression, $\hat{\beta}'_{KR}$ and $\hat{\beta}'_{LR}$ are the estimation of the slope of Security Market Line (Curve). For the parameters of linear regression, confidence is indicated by ** and * at confidence level 99% and 95%, respectively.

II.2 Multi-factor models

We introduce the multivariate extension of univariate non-parametric methodology for testing the linearity of multi-factor models (Fama-French three-factor model and Carhart four-factor model). The generalized multivariate formula of Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964) is

$$\hat{y} = \hat{m}_{\mathbf{H}}(\mathbf{x}) = \sum_{i=1}^n W_{\mathbf{H}_i}(\mathbf{x}) y_i, \quad (\text{II.8})$$

where multivariate Nadaraya-Watson weighting is

$$W_{\mathbf{H}_i}(\mathbf{x}) = \frac{\kappa_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x})}{\sum_{i=1}^n \kappa_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x})}, \quad (\text{II.9})$$

and $\kappa_{\mathbf{H}}$ is the multivariate kernel function using \mathbf{H} bandwidth matrix and $\mathbf{x} = (x_1, x_2, \dots, x_d)$ is a d -dimensional explanatory vector. In multivariate setting, the goal is to select the optimal multivariate kernel function and bandwidth matrix. As Fama and French (1993) argue that the market proxy, the SMB and HML factors exhibit non-significant correlation, we simplify multivariate kernel function by using a diagonal bandwidth matrix \mathbf{H} with one bandwidth per variable, resulting

$$\kappa_{\mathbf{H}}(\mathbf{u}) = \frac{\kappa(\mathbf{H}^{-1}\mathbf{u}^T)}{\det(\mathbf{H})} = \prod_{k=1}^d \frac{1}{h_k} K\left(\frac{u_k}{h_k}\right). \quad (\text{II.10})$$

Based on that, we reduce the optimization problem to d univariate selection problem that has already been discussed. We apply multivariate R^2 for measuring the accuracy of regression models. The multivariate hypothesis testing is similar to univariate method; in this case, we use

$$\text{multivariate operator with the error function } T_M = \sum_{i=1}^n \left(\hat{m}_{\mathbf{H}}(\mathbf{x}_i) - \sum_{j=1}^n W_{\mathbf{H}_i}(\mathbf{x}_j) m_{\hat{\theta}}(\mathbf{x}_j) \right)^2.$$

We introduce non-parametric estimation of coefficients of multi-factor models. We approximate the derivatives of $m(\mathbf{x})$ in $\mathbf{x} = (x_1, x_2, \dots, x_d)$ by kernel-weighted least squares

$$\hat{\boldsymbol{\beta}}(\mathbf{x})^T = \left(\hat{\boldsymbol{\beta}}^{(0)}(\mathbf{x}), \hat{\boldsymbol{\beta}}^{(1)}(\mathbf{x}), \dots, \hat{\boldsymbol{\beta}}^{(p)}(\mathbf{x}) \right)^T = (\mathbf{D}^T \mathbf{W} \mathbf{D})^{-1} \mathbf{D}^T \mathbf{W} \mathbf{y}, \quad (\text{II.11})$$

where $\hat{\boldsymbol{\beta}}^{(j)}(\mathbf{x})$ is the j^{th} derivative estimation vector of $m(\mathbf{x})$, \mathbf{D} is an $n \times (pd + 1)$ matrix generated from the observation of explanatory variables in order $j=0, 1, \dots, p$, \mathbf{W} is a weight matrix and \mathbf{y} is the observation of dependent variable. The non-parametric coefficients is estimated by

$$\hat{\boldsymbol{\beta}}_{KR} = \frac{1}{n} \sum_{i=1}^n \hat{\boldsymbol{\beta}}^{(1)}(\mathbf{x}_i), \quad (\text{II.12})$$

and the estimation of non-parametric alpha is the following:

$$\hat{\alpha}_{KR} = \frac{1}{n} \sum_{i=1}^n y_i - \hat{\boldsymbol{\beta}}_{KR} \mathbf{x}_i^T. \quad (\text{II.13})$$

In our multi-factor analysis, we apply linearity testing of multi-factor models, furthermore non-parametric estimation of coefficients and abnormal performance for all investigated stocks. Based on the hypothesis testing, the lowest value of p is 0.08 for Fama-

French model, and 0.09 for Carhart model, thus the linearity cannot be rejected for neither Fama-French three-factor model nor Carhart four-factor model at confidence level 95% and the linear estimation of coefficients is adequate. Table II.4 summarizes the average results aggregated by size of market capitalization (S&P 500, S&P MidCap 400 and S&P SmallCap 600) and for all companies using Fama-French three-factor model (Panel A) and Carhart four-factor model (Panel B). Although the linearity cannot be rejected, the kernel regression estimation shows significantly better goodness of fit. For both models, we find that the coefficient of HML ($\hat{\beta}_{H,LR}$) is significantly overestimated by linear regression.

Table II.4. Summary of alpha- and coefficient estimation of multi-factor models

Panel A – Fama-French three-factor model

Segment	$\overline{R_{KR}^2}$	$\overline{R_{LR}^2}$	$\overline{\hat{\alpha}_{KR}}$	$\overline{\hat{\alpha}_{LR}}$	$\overline{\hat{\beta}_{3,KR}}$	$\overline{\hat{\beta}_{3,LR}}$	$\overline{\hat{\beta}_{S,KR}}$	$\overline{\hat{\beta}_{S,LR}}$	$\overline{\hat{\beta}_{H,KR}}$	$\overline{\hat{\beta}_{H,LR}}$
S&P 500	0.38	0.28	0.034	0.034	1.041	1.025	0.190	0.107	0.200	0.249
S&P MidCap 400	0.36	0.27	0.036	0.034	1.024	1.034	0.591	0.571	0.393	0.508
S&P SmallCap 600	0.32	0.22	0.046	0.043	0.965	0.971	0.855	0.861	0.226	0.371
All companies	0.35	0.26	0.039	0.037	1.010	1.010	0.545	0.513	0.273	0.376

Panel B – Carhart four-factor model

Segment	$\overline{R_{KR}^2}$	$\overline{R_{LR}^2}$	$\overline{\hat{\alpha}_{KR}}$	$\overline{\hat{\alpha}_{LR}}$	$\overline{\hat{\beta}_{4,KR}}$	$\overline{\hat{\beta}_{4,LR}}$	$\overline{\hat{\beta}_{S,KR}}$	$\overline{\hat{\beta}_{S,LR}}$	$\overline{\hat{\beta}_{H,KR}}$	$\overline{\hat{\beta}_{H,LR}}$	$\overline{\hat{\beta}_{M,KR}}$	$\overline{\hat{\beta}_{M,LR}}$
S&P 500	0.47	0.29	0.037	0.038	1.00	1.00	0.21	0.12	0.17	0.22	-0.059	-0.079
S&P MC. 400	0.46	0.28	0.041	0.041	0.97	0.99	0.59	0.59	0.34	0.45	-0.118	-0.156
S&P SC. 600	0.40	0.22	0.051	0.048	0.94	0.94	0.87	0.88	0.23	0.32	-0.161	-0.132
All companies	0.44	0.26	0.043	0.043	0.97	0.98	0.55	0.53	0.25	0.33	-0.113	-0.122

Notes: In the table, we show the average goodness of fit and coefficients of the Fama-French three-factor model (Panel A) and Carhart four-factor model (Panel B) of 50-50 randomly chosen companies from the S&P 500, S&P MidCap 400 and the S&P SmallCap 600 universe. We use cross-sectional kernel- and linear regression between the series of risk premium of stocks and risk premium of market portfolio and risk factors. For Fama-French model we apply SMB and HML factors, for Carhart model we extend the factors by MOM factor. SMB is the average return between stocks of small and large companies; HML is the average return between high book-to-market ratios minus low ones; MOM is momentum that is an empirically observed tendency for rising asset prices to rise further and falling prices to keep falling. For kernel regression, we use cross-validation-based optimally selected bandwidth matrix, multivariate Gaussian kernel function and the Nadaraya-Watson weighting function. The 1st column is the label of segment in we aggregate the results. The 2nd and 3rd column shows the average goodness of fit regression models, the following columns summarizes the average alpha, beta and coefficients of SMB, HML and momentum factors by kernel- and linear regression, alternately.

III Entropy-based Asset Pricing

We investigate entropy as a risk measure in capital market. For asset pricing, we define differential Shannon- and Rényi entropy of risk premiums of an asset as an alternative measure of risk; furthermore, we discuss estimation methods for differential entropy. We investigate whether the entropy-based risk measures satisfy the axioms of coherent risk. We introduce an evaluation methodology that is capable to compare in-sample explanatory and out-of-sample predictive power of various risk measures. For baseline measures of evaluation, we use standard deviation and CAPM beta. We also measure how entropy behaves in the function of securities involved in a portfolio. As an outlook, we extend our methodology for multivariate risk models and investigate the capabilities of entropy-based risk measures for extending multi-factor models. For empirical investigation, we use daily logarithmic returns of risk free asset and 150 randomly selected stocks from S&P 500 index for a 27-year period; furthermore, we apply the proxy of market portfolio for beta, and daily observations of SMB, HML and momentum factors for multi-factor models.

III.1 Entropy as risk

We consider differential entropy as a measure of uncertainty of a random continuous variable X . We apply two special cases of generalized entropy function (Rényi, 1961)

$$H_{\alpha}(X) = \frac{1}{1-\alpha} \ln \int f(x)^{\alpha} dx, \quad (\text{III.1})$$

which is Shannon entropy ($\alpha = 1$)

$$H_1(X) = -\int f(x) \ln f(x) dx \quad (\text{III.2})$$

and Rényi entropy ($\alpha = 2$)

$$H_2(X) = -\ln \int f(x)^2 dx. \quad (\text{III.3})$$

The “plug-in” estimation of generalized entropy function is

$$H_{\alpha,n}(X) = \frac{1}{1-\alpha} \ln \int_{A_n} f_n(x)^{\alpha} dx, \quad (\text{III.4})$$

where $f_n(x)$ is the estimation of probability density function of based on the observation of X ,

and A_n is the range of integral. Based on the investigation of most well-known probability density estimation functions, we find that histogram-based estimation offers the best tradeoff between the in-sample explanatory and out-of-sample predictive power. The estimation of $f(x)$ by histogram is $f_n(x) = \frac{v_j}{nh}$, if $x \in (t_j, t_{j+1})$, where v_j is the number of data points falling in the j^{th} bin, h is the size of a bin. As the plug-in estimation contains integral operator that makes the implementation of the estimator difficult, we deduce a simpler “built-in” estimator formula for Shannon- and Rényi entropy:

$$H_{1,n}(X) = -\frac{1}{n} \sum_{j=1}^g v_j \ln \left(\frac{v_j}{nh} \right), \quad (\text{III.5})$$

$$H_{2,n}(X) = -\ln \sum_{j=1}^g h \left(\frac{v_j}{nh} \right)^2 \quad (\text{III.6})$$

As the differential entropy $H_\alpha(X)$ is not positive homogeneous and it can take negative values, we apply a transformation with exponential function thus we define our entropy-based risk measure κ for asset A by the following formula:

$$\kappa_{H_\alpha}(A) = e^{H_\alpha(R_A - R_F)}, \quad (\text{III.7})$$

where R_A is the random variable for return of asset A and R_F is the risk-free rate. We see that it captures risk without using any information about the market. We show that if $R_A - R_F$ follows normal distribution, the Shannon entropy-based risk measure differs from standard deviation in a constant $\sqrt{2\pi e}$ only.

Based on Artzner et al. (1999) a risk measure is considered as a coherent measure if it satisfies the axioms of translation invariance, subadditivity, positive homogeneity and monotonicity. In this dissertation, we show that entropy-based risk measure defined by Eq. (III.7) is positive homogeneous; furthermore it is subadditive and convex if the distribution of the return of portfolios is normal. We analytically deduce that entropy-based risk measure does not satisfy the axioms of translational invariance and monotonicity; therefore, it is not coherent. However, we note that coherence is not a requirement for asset pricing and we show that entropy can be used efficiently for explaining and predicting expected return of assets in capital market.

III.2 Empirical results

Based on random generation of 1 million pairs of randomly weighted portfolios, we empirically find that entropy is capable to capture the diversification effect for any two portfolios with 99% confidence. In another measurement, by generating 10 million equally weighted portfolios, we measure their risk by entropy-based risk measures and standard deviation and we calculate the averages based on the number of elements involved. We illustrate the results in Figure III.1. For 10 random securities involved in the portfolio, approximately 40% of risk reduction can be achieved compared to a single random security, based on all of the three investigated risk estimators. The figure suggests that entropy behaves similar but not the same as standard deviation.

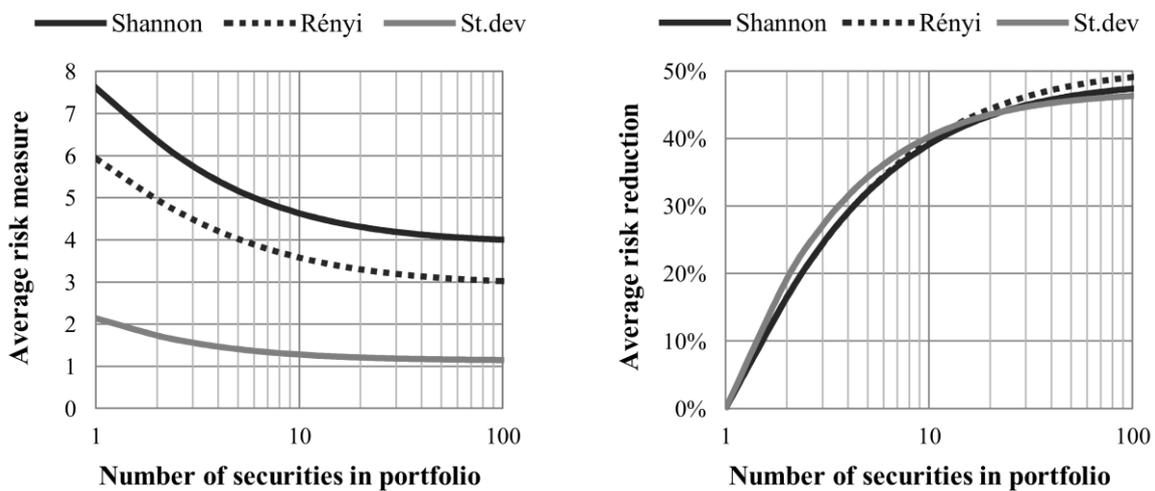


Figure III.1. Average value of risk and risk reduction vs. number of securities in portfolio

Notes: We generate 10 million random equally weighted portfolios with various number of securities involved (at most 100,000 for each size) based on 150 randomly selected securities from S&P 500. The risk of portfolios is estimated by standard deviation (gray continuous curve), Shannon- (black continuous curve) and Rényi entropy (black dashed curve) in the period from 1985 to the end of 2011 and averaged by the number of securities involved in a portfolio. The left chart shows the average risk estimates for each portfolio size, and the right chart shows the risk reduction compared to an average risk of a single security.

We also investigate how the various portfolios behave in the expected return – risk coordinate system in the function of diversification. We generate 200-200 random equally weighted portfolios with 2, 5 and 10 securities involved, and compare these to single securities using Shannon- and Rényi entropy. Figure III.2 shows that random portfolios are situated on a hyperbola as in the portfolio theory of Markowitz (1952).

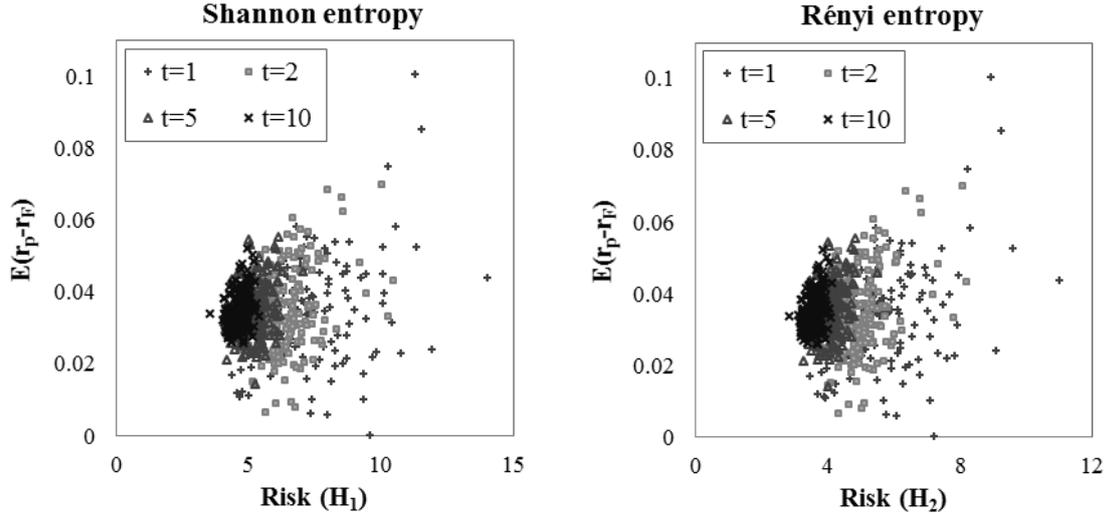


Figure III.2. Random portfolios with various number of securities involved in expected risk premium – risk coordinate system

Notes: The panels show the expected risk premium of the portfolios (calculated by the average of daily risk premiums) versus the estimated risk using various methods; the number of securities involved (t) is indicated by different markers. We generate a sample of 750 random portfolios by using 150 randomly selected securities and 200-200 random equally weighted portfolios with 2, 5 and 10 securities. The risk of portfolios is estimated by Shannon- and Rényi entropy by using daily returns in the period from 1985 to the end of 2011.

We introduce a method to compare the explanatory and predictive power of any single-variate risk measure. We consider the estimated risk measures as explanatory variables of the expected risk premium in the same period (in-sample explanation) or in the next period (out-of-sample prediction). We define explanatory- and predictive power as the goodness of fitting (R^2) of the regression model of in-sample and out-of-sample settings. By defining various kind of samples, we evaluate Shannon- and Rényi entropy against standard baselines as standard deviation and CAPM beta. The results of measurement and the description of samples are summarized in Table III.1.

Table III.1. Comparison of accuracy of risk measures in various samples

Risk measure	$R_{I,lt}^2$	$R_{I,bull}^2$	$R_{I,bear}^2$	$\overline{R_{I,st}^2}$	$\overline{R_{O,st}^2}$	$\sigma_R(R_{I,st}^2)$	$\sigma_R(R_{O,st}^2)$
Standard deviation	0.0783	0.3390	0.3671	0.0794	0.0970	0.75	0.65
CAPM beta	0.0617	0.3667	0.4369	0.1331	0.0645	0.98	1.02
Shannon entropy	0.1298	0.4345	0.3961	0.1338	0.1015	0.69	0.64
Rényi entropy	0.1571	0.4236	0.3855	0.1282	0.0934	0.62	0.60

Notes: The table summarizes the explanatory- (in-sample R^2) and predictive power (out-of-sample R^2) of the investigated risk measures in various samples. We estimate risk measures of 150 random securities using standard deviation, CAPM beta, Shannon- and Rényi entropy risk estimation methods for (1) long term, from 1985 to the end of 2011 (1985-2011); (2) long term on upward trends (bull market), (3) long term on downward trends (bear market), (4) 18 10-year periods shifting by one year from period (1985-1994) to period (2002-2011), split into two 5-5 year periods for each. The 2nd column shows the explanatory power of risk measures for long term. The 3rd

and 4th column are the explanatory power on upward and downward trends, respectively. The 5th column stands for the average explanatory power of risk measured in the first 5 years of 10-year shorter periods in sample. The 6th stands for the average predictive power of risk measures (out-of-sample R^2) calculated by estimating risk in the first 5 years and evaluating them on the other 5 years in each 10-year periods. The last two columns show the relative standard deviation of explanatory and predictive power of short-term samples.

Based on Table III.1, we state the followings. Shannon entropy beats standard deviation and CAPM beta in long- and short term; furthermore, the Rényi entropy has the best long-term explanatory power. Both Shannon- and Rényi entropies give more reliable risk estimation; their explanatory power exhibits significantly lower variance compared to CAPM beta. The reliability of standard deviation and entropies is statistically not different. If upward and downward trends are distinguished, the regime dependency of entropy is recognized and higher explanatory power is measured for all risk measures; however, the relation between the accuracy of entropy and CAPM beta exhibits mixed results.

As an extension of single factor risk measures, we investigate multivariate risk models. We find that Fama-French- and Carhart models have significantly better explanatory and predictive power. We show that entropy is capable to improve the accuracy of these models; however, the performance gain decreases by diversification. We find that combining higher moments can also improve the explanatory and predictive power of entropy. In the dissertation, we discuss several other combination of multivariate risk measures.

References

- 1 Artzner, P., Delbaen, F., Eber, J. & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3), 203—228. doi:10.1111/1467-9965.00068
- 2 Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1), 3—18. doi:10.1016/0304-405X(81)90018-0
- 3 Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of financial economics*, 12(1), 129—156. doi:10.1016/0304-405X(83)90031-4
- 4 Clausius, R. (1870). XVI. On a mechanical theorem applicable to heat. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 40(265), 122—127. doi:10.1080/14786447008640370
- 5 Fama, E. F. & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427—465. doi: 10.1111/j.1540-6261.1992.tb04398.x
- 6 Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56. doi:10.1016/0304-405X(93)90023-5
- 7 Fama, E. F. & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1), 55—84. doi:10.2307/2329302
- 8 Härdle, W., Müller, M., Sperlich, S., Werwatz, A. (2004): *Nonparametric and Semiparametric Models*. Springer Series in Statistics, Springer-Verlag. Chapter 1-4. doi:10.1007/978-3-642-17146-8
- 9 Lintner, J. (1965a). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, 47(1), 13—37. doi:10.2307/1924119
- 10 Lintner, J. (1965b). Security Prices, Risk and Maximal Gains from Diversification. *Journal of Finance*, 20(4), 587—615. doi:10.1111/j.1540-6261.1965.tb02930.x
- 11 Markowitz, H. (1952). Portfolio selection*. *The Journal of Finance*, 7(1), 77—91. doi:10.1111/j.1540-6261.1952.tb01525.x
- 12 Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768—783. doi:10.2307/1910098
- 13 Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1), 141—142. doi:10.1137/1109020
- 14 Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, 19(3), 425—442. doi:10.2307/2977928
- 15 Treynor, J. L. (1962). Toward a Theory of Market Value of Risky Assets. Robert Korajczyk (Ed.), *Asset Pricing and Portfolio Performance*. London: Risk Books. 1999
- 16 Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4), 359—372.

Related publications by the author

Referred by the theses (Journal Papers)

- 1 Erdős, P., Ormos, M. & Zibriczky, D. (2011). Non-parametric and semi-parametric asset pricing. *Economic Modelling*, 28(3), 1150—1162. doi:10.1016/j.econmod.2010.12.008 (IF: 0.701)
- 2 Ormos, M. & Zibriczky, D. (2014). Entropy-Based Financial Asset Pricing. *PLoS ONE*, 9(12): e115742. doi:10.1371/journal.pone.0115742 (IF: 3.234)
- 3 Erdős, P., Ormos, M. & Zibriczky, D. (2010b). Non-parametric Asset Pricing: Evidence from US Stocks. *The Empirical Economics Letters*, 9(6), 573—580
- 4 Erdős, P., Ormos, M. & Zibriczky, D. (2010a). Egyenes-e a tőkepiaci árazási modell (CAPM) karakterisztikus és értékpapír-piaci egyenese?. *Közgazdasági Szemle*, 57(3), 201—221.

Journal Papers

- 5 Ormos, M. & Zibriczky, D. Entrópia mint pénzügyi kockázati mérték. *Sigma*, Manuscript submitted for publication, revise and resubmit.

Conference Papers

- 6 Ormos, M. & Zibriczky D. (2013, June). Entropy Based Asset Pricing. In Aaro Hazak (Ed.), *5th International Conference “Economic Challenges in Enlarged Europe”*. Tallinn, Estonia (pp. 1—20, paper 2).
- 7 Ormos, M. & Zibriczky, D. (2013, June). Asset pricing and entropy. In – *Proceedings of the 10th International Scientific Conference on European Financial Systems 2013*, Brno, Czech Republic, (pp. 241—248).
- 8 Erdős, P., Ormos, M. & Zibriczky, D. (2010, May). Non-linear Asset Pricing. In L. Galetic, M. Spremic. & M. Ivanov (Eds.), *5th International Conference “An Enterprise Odyssey: From Crisis to Prosperity—Challenges for Government and Business”*. Opatija, Croatia (pp. 564—590). (ISBN: 953-6025-34-5)
- 9 Erdős, P., Ormos, M. & Zibriczky, D. (2010, January) Kernel Based Asset Pricing. In: P. Cervinek & P. Musil (Eds.), *2nd international PhD students conference – New Economic Challenges*. Brno, Czech Republic (pp. 16—23). (ISBN:978-80-210-5111-9)

Posters

- 10 Erdős, P., Ormos, M. & Zibriczky, D. (2009, June). *Non-parametric Asset Pricing: Evidence from U.S. Stocks*. Poster presented at Morgan Stanley-BME Financial Innovation Centre Kick-off & Workshop, Budapest, Hungary