Management Studies, Sep.-Oct. 2022, Vol. 10, No. 5, 294-324
doi: 10.17265/2328-2185/2022.05.004

Some Financial Problems in the Light of EMM Results: Asset Pricing and Efficient Portfolio Allocation

Valery V. Shemetov

Discussing results in asset pricing and efficient portfolio allocation, we show that mixed success and errors in these results often follow from a lack of information about the asset return distribution and wrong assumptions about its properties. Some mistakes in asset pricing come from the assumption of symmetry in return distributions. Some errors in efficient portfolio allocation follow from Markowitz's approach when applying it to portfolio optimization of skewed asset returns. The Extended Merton model (EMM), generating skewed return distributions, demonstrates that (i) in skewed asset returns, the variance is not an adequate measure of risks and (ii) positive skewness in the asset returns comes together with a high default probability. Thus, the maximization of the mean portfolio returns and skewness with controlled variance used in mainstream papers can critically increase portfolio risks. We present the new settings of the optimal portfolio allocation problem leading to less risky efficient portfolios than the solutions suggested in all previous papers.

Keywords: asset pricing, efficient portfolio allocation, skewed returns, default probability, Extended Merton model

Introduction and EMM Description

The seminal article (Merton, 1974) has determined the development of financial economics for a long period and continues its influence in many sectors of the financial theory. The article considers the dynamics of firm's asset returns and applies the general theory to estimating the value of a zero-coupon bond. Strengthened with some important extents, Merton's model (another name is the geometric Brownian model, GBM) has become very popular among the financial theorists investigating credit risk problems (e.g., Cox, Ross, & Rubinstein, 1979; Harrison & Pliska, 1981; Leland, 1994; Hugonnier, Malamud, & Morellec, 2015). Recently, an Extended Merton model (EMM) has appeared (Shemetov, 2020a) that sheds more light on the development of asset returns of the firm. EMM solves the general problem of asset returns by analyzing the firm dynamics with the effects of payments. The principal difference between GBM and EMM solutions is that GBM produces a normal distribution for the asset returns while EMM proves that the return distribution starting with the normal distribution becomes more and more skewed. Our goal is to expose typical errors in asset pricing models and solutions to the optimal portfolio allocation problem following from the prejudice that GBM produces reliable return distributions. To fulfill the program, we review the papers from these fields in the theory of investments and explain why these papers meet mixed success in their attempts to solve the problems. Section 1 includes the introduction and a short description of EMM, Section 2 considers typical papers on asset

Acknowledgment: I am infinitely thankful to my friend and colleague M. Rubinstein for valuable discussions, deep and constructive remarks, and an invariable interest in my work. Of course, all errors are my full responsibility.

pricing, Section 3 discusses the papers on the optimal portfolio allocation problem, Section 4 presents results of EMM modeling of the asset return distribution, and Section 5 concludes the paper.

Here we briefly introduce the reader to EMM (Shemetov, 2020a; 2020b; 2021), which we extensively use in this paper. We consider asset returns in a logarithmic form $x = \ln(RX/P_0)$, R is the rate of expected returns, $X(t)$—firm's stochastic assets, $P > 0$—business securing expenses (BSEs), $P = FC + DP + TAX + DIV$, *FC*—fixed costs, *DP*—debt payments, *TAX*—taxes, *DIV*—dividends, all dollar per time unit; $P(t) = P_0\pi(t)$, $\pi(t)$—a regular piecewise continuous function of time. For $x > 0$, the assets grow, for $x < 0$, they decline, $x(t) = 0$ is a line of unstable equilibrium. The return distribution $V(x,t)$ meets the equation (V_y is a partial derivative over variable y).

$$V_t + R(1 - \pi(t)e^{-x})V_x - 0.5C^2V_{xx} + R\pi(t)e^{-x}V = 0 \tag{1}$$

starting its evolution from a normal distribution (the initial condition):

$$V(x,0) = N(x; H_0, \sigma_0^2),\ H_0 = \langle x(0)\rangle = \langle \ln[RX(0)/P_0]\rangle,\ \sigma_0^2 = \langle (x(0) - H_0)^2\rangle \tag{2}$$

C^2 is the intensity of random market shocks. To solve the problem, one must add a boundary condition for the return distribution, reflecting the fact that when the firm's assets intercept the boundary, the firm defaults:

$$DL = \max\{\ln(RX_D/P_0), 0\} \tag{3}$$

The line $DL = \ln(RX_D/P_0)$ corresponds to the outstanding debt X_D (the exogenous default line). The line $DL = 0$ serves as a soft default line (the endogenous default line) because when $x < 0$, business activities inflict losses to the firm. If the problem settings do not include the default line (see Merton, 1974; the marginal case when the firm's assets are extremely large, tend to the infinity), the return distribution starting from the normal distribution remains normal, spreading with the rate C. In real cases with default line DL, the mean value will rise or go down depending on the difference $RX - P$, and the return distribution decreases with rate $IDP(t)$, the intensity of default probability:

$$IDP(t) = 2\int_{-\infty}^{DL(t)} V(x,t)dx \tag{4}$$

and default probability $DPR(t)$ over the interval $[0, t]$, t—the time of observation, is

$$DPR(t) = \int_0^t IDP(\tau)d\tau \tag{5}$$

The rate of distribution leak through the boundary is proportional to the part of distribution $V(x, t)$ below the default line. The loss of mass (Brownian particles constituting the distribution) at the default line induces redistribution of the remaining mass down to the default line. The distribution becomes negatively skewed. The cumulative effect of the diffusion spread and the distribution distortion makes the negative tail grow heavier over time increasing the default probability. See details in Shemetov (2020a).

So, the negatively skewed distribution $V(x, t)$ of asset returns is a *natural effect* of the firm dynamics. This fact makes void the calibrated models representing a combination of GBM with empirical default databases (e.g., Huang & Huang, 2003; Bohn, 2006). See a discussion of calibrated models below on the example of Moody's KMV. EMM also constrains application of the jump-diffusion processes (e.g., Zhou, 2001; Hilberink & Rogers, 2002; Kou, 2002; Chen & Kou, 2009), which join a standard diffusion process with poorly motivated jumps down in asset returns. The objective of adding jumps to the diffusion is to raise the default probabilities over the low values specific for GBM and make theoretical default probabilities comparable with the default frequencies observed in practice.

The fact that asset returns are skewed is now well established in the optimal portfolio allocation theory and the asset pricing theory. However, the nature and driving forces of asset skewness remain still unknown to the economists and investors unfamiliar with EMM. The novel vision of the asset return development helps to improve the comprehension of some traditional problems of financial management. EMM used for analysis of the effects of firm leverage (debt) demonstrates that debt *negatively* affects the returns and survival of the firm. EMM reveals inconsistency in the debt study (Leland, 1994) that makes a basis for the tradeoff theory. This theory claims the existence of the optimal leverage maximizing the firm returns and strengthening its economic state. One can find a critical discussion of Leland (1994) and the tradeoff theory in Shemetov (2021).

So far, EMM is the only model that can analyze the *effects of inflation* on the firm returns and survival (Shemetov, 2020b). It seems to be the very time to take account of the current inflation growing worldwide. EMM proves that moderate inflation acts positively on the firm's returns and survival, while high inflation depresses the returns and decreases the firm chances for survival. The paper shows that the threshold separating moderate inflation from high inflation depends on the state of the firm: the same level of inflation can be mild for one firm and high for another. It means that there is no general recipe for the firm to deal with the effects of inflation; this reaction must be strictly individual.

Because of the continuous leak of Brownian particles at the default line in EMM, the mean of the stochastic part of asset returns is always negative, and the stochastic returns never make a martingale. When the firm is far from default, the leak is low making an illusion that the asset returns are a martingale. If the firm comes closer to default, the leak grows noticeably, and asset returns cannot be assumed a martingale anymore. That is why the no-arbitrage principle is effective in the market of GBM firms (no leak at the default line) and ineffective in the market of EMM firms. So, all theoretical constructions admitting no-arbitraging have a limited domain of their validity and are not unconditionally correct. Inefficiency of the no-arbitraging principle in the market of the firms paying BSEs cancels the *risk-neutral approach* and *risk-neutral probabilities* (e.g., Cox et al., 1979; Harrison & Kreps, 1979; Harrison & Pliska, 1981) in the analysis of credit risks.

In Markovitz's mean-variance portfolio optimization, the investor maximizes the mean portfolio return, keeping the variance fixed. At that, the variance is considered a good measure of risk, which is true for symmetric distributions. Considering optimal portfolio allocation with higher moments (skewness, kurtosis), Kraus and Litzenberger (1976) recommend selecting positively skewed assets. So, the researchers looking for the efficient portfolio with higher moments maximize the portfolio mean return and skewness while keeping the variance fixed. We show in Section 4 that high positive skewness comes with a high default probability, and the unconditional maximization of skewness dangerously increases portfolio risks. We also show that for skewed asset returns, the variance is no adequate measure of asset risks and its minimization becomes useless. For skewed return distributions, the default probability, or the intensity of default probability, is the proper measure of asset risks. In the portfolio optimization with higher moments, one must solve the following problem: *maximize the mean return and skewness while keeping the default probability under control*. Such approach helps select both efficient and reliable portfolios with controllable risks.

To compute the observable statistical moments of the return distribution like the mean return, variance, skewness, etc., one must consider the conditional return distribution including the firms with the same prehistory and existing at the time of observation t. One can get this distribution $\hat{V}(x,t)$ from distribution $V(x,t)$ as

$$\hat{V}(x,t) = V(x,t) - V(2DL - x, t) \qquad (6)$$

determined in the interval $[DL, +\infty)$ and having the property $\hat{V}(DL,t) = 0, t \in [0,+\infty)$. The moments of the return distribution $\hat{V}(x,t)$ are

$$\hat{H}(t) = \int_{DL}^{\infty} x\hat{V}(x,t)dx \ , \ \widehat{Var}(t) = \int_{DL}^{\infty} (x-\hat{H})^2 \hat{V}(x,t)dx \ ,$$
$$\widehat{SK}(t) = \int_{DL}^{\infty} (x-\hat{H})^3 \hat{V}(x,t)dx \ , \ \widehat{MX}(t) = (P_0/R)\int_{DL}^{\infty} e^x \hat{V}(x,t)dx \qquad (7)$$

They are denoted in Section 4 as *H_BP*(*t*), *VAR_BP*(*t*), *SK_BP*(*t*), and *MX_BP*(*t*), correspondently (BP stands for "Boundary Problem"). The parameters governing the evolution of return distribution are: the distance to default $H(t) - DL(t)$, variance $\sigma^2(t)$, skewness *SK*(*t*), the intensity of market shocks C^2, and the stochastic ratio $RX(t)/P(t)$. One can add here a general economic motion influencing the firm development: a growing economy with higher rates of returns and lesser default risks and a declining economy with lower rates of returns and higher default risks.

Reading the review below, one must remember that many authors imagine the return distribution as symmetric (normal or, more generally, elliptic functions). The statistical characteristics they operate with depend only on the mean and variance of returns. We know now that the distribution of asset returns is skewed and the mean and variance do not provide for exhaustive information on the firm. Therefore, all estimates presented in those papers are conditionally correct, as far as the return distribution remains almost symmetric (GBM).

We do not plan to present a comprehensive review of numerous attempts to solve the problems of asset pricing or portfolio optimization. Our objective is much humbler: to show the dramatism of the scientific search for a better CAPM and optimal portfolio allocation and give some new evidence casting more light on these problems. We hope that this information will be helpful in the nearest future.

Asset Pricing Problem

Portfolio optimization is most important for the theory and practice of investments. Being in the center of the vivid discussion since 1952, when Markowitz has published his seminal mean-variance analysis of efficient portfolios, this theory is still far from its satisfactory resolution despite all progress achieved in this area. The Capital Asset Pricing Model, CAPM (Sharpe, 1964; Lintner, 1965; Mossin, 1966), developing Markowitz's ideas, relates the expected portfolio return $E(r_p)$ with the expected market return $E(r_M)$ and beta:

$$E(r_p) = r_f + \beta_p[E(r_M) - r_f] \ , \ \beta_p = cov(r_p, r_M)/var(r_M) \qquad (8)$$

here r_f is a return on a risk-free asset. CAPM uses two basic assumptions (i) asset returns are normally distributed or (ii) investors have a quadratic utility function. Notwithstanding these rigid limitations, practitioners widely use CAPM for asset allocation because it provides an intuitively clear assessment of the relative merits of alternative portfolios. Moreover, one can span CAPM mean-variance frontiers with only two funds, which makes it easy to calculate and interpret these frontiers. Financial theorists (Merton, 1974; Leland, 1994; and many others) use CAPM in their analysis because it is consistent with GBM, which now dominates the other methods in credit risk analysis. Sentana (2009) gives a comprehensive survey on the mean-variance efficiency tests. We demonstrate below that using old good Markowitz's portfolio allocation is less dangerous for the investors than following recommendations of the unripe mean-variance-skew portfolio allocation theories published before the summer of 2022.

Empirical and theoretical studies show that both asset and portfolio returns are not normally distributed. Brown and Matysiak (2000) show that the return distribution has heavier tails than one can expect for a normal distribution. Simkowitz and Beedies (1980), Singleton and Wingender (1986), Badrinath and Chatterjee (1988), and Harvey and Siddique (2000) present empirical evidence of skewness in individual stock returns as well as market indexes in the US stock markets. Shemetov (2020a; 2020b; 2021) demonstrates that an initially normal distribution of asset returns becomes skewed over time.

Theoretical studies of Kraus and Litzenberger (1976) start exploring the role of skewness in the behavior of asset returns. They suppose that the utility function is not quadratic but admits the third moment (skewness) characterizing the return distribution asymmetry. The authors argue that investors will value the positive skewness. When two assets have the same expected returns and variances, the investors consider these two assets equivalent unless they know the asset skewness values. The authors insist that in this case, the wealth-maximizing investors will prefer the asset with the highest possible skewness increasing their returns the most. Kraus and Litzenberger (1976) consider one desirable side of the optimal portfolio (its high returns) and pay no attention to another important side of the optimal portfolio, its stability. As we show in Section 4, their recommendation to choose assets with the highest possible skewness leads to *uncontrollable growth of portfolio risks threatening the investor with sudden losses*. Kraus and Litzenberger (1976) start the stream of portfolio optimization studies developing the mean-variance-skewness analysis (the MVS analysis). Some researchers (e.g., Fang & Lai, 1997; Dittmar, 2002; Jondeau & Rockinger, 2003; Krueger, 2021) go further and include the fourth moment (kurtosis) in their MVSK analysis, but there are a lot of vague moments even in the MVS analysis.

We distinguish three streams in the literature considering the asset pricing and portfolio optimization problems. (i) The papers using heuristic/fundamental variables instead of or together with the beta to predict expected asset returns. (ii) The MVS investigations where the skewness of asset returns appears through proxies. (iii) The studies modeling asymmetry of the return distribution with jump-diffusion processes or special (skew-elliptic) functions that help solve the portfolio optimization problem in closed form. Many of these studies use regression analysis of selected variables on the expected asset returns. Their logic is like this. To solve the optimal portfolio problem, investors must have an asset pricing model relating portfolio payoffs with market factors. Economists see this model as a linear valuation functional that assigns prices to market factors (Hansen & Jagannathan, 1997). Because the asset pricing model is unknown and market factors are often unobservable, economists have to construct approximating models using empirical proxies for the theoretical market factors. As a result, empirical models deviate from the objective asset pricing models due to possible misspecification. To find the best among the available empirical models, one must know how to compare the competing models. Shanken (1987) and Hansen and Jagannathan (1997) suggest solutions to this problem by proposing the measures of model misspecification that are widely used for parameter estimation, specification testing, and comparison of the competing empirical models.

We now introduce some factors that appear in many asset-pricing models. *The firm size* and *book-to-market* ratio are among such factors. According to EMM, a small firm has characteristics: low mean assets and, therefore, a short distance to default. A probable lack of management skills increases the variance of returns to higher values. These circumstances make the negative tail of the return distribution grow fast (high skewness), threatening the firm with soon default (high default risks). To survive, the firm takes riskier projects (high *C*), promising higher expected rates of return, *R*. Some small firms succeed in such projects, increasing their distance to default and achieving the state of a medium-sized firm. Struggling for survival and success, these small firms show returns higher than large firms. At that, the variance of returns of a small firm is not much greater than that of a large firm because, for asymmetric distributions, the variance is a not-so-good measure of the default risk, as it is for symmetric distributions. But many small firms default, failing to bring their risky projects to success. It explains the high mortality of small firms. In contrast, a large firm has high mean assets, a long distance to default, and a low default probability, all other parameters equal. Such firms, as a rule, value their stability and avoid risky projects preferring to have a moderate but steady rate of average returns. Summarizing: the small firm has higher skewness and a greater default risk than a large firm (all other parameters equal). However, the skewness and default risk are not one-to-one functions of the firm size; researchers must use this variable cautiously (see Section 4). A high *book-to-market* ratio of the firm means that the market estimates its perspectives skeptically, and its income and dividends are low. It can happen to a firm with a short distance to default that suffers financial distress or slowly recovers from it. In both cases, the return distribution is noticeably skewed, threatening the firm with likely default. A low book-to-market value means that the market estimates the firm default probability as insignificant and its business prospects favorable. Again, the skewness and default risk are not one-to-one functions of the book-to-market ratio.

Concerning the “fundamental” variables (e.g., Daniel & Titman, 1997; Shapovalova, Subbotin, & Chauveau, 2011; Harvey, Liu, & Zhu, 2016), one can see that they are not among the parameters and variables governing the evolution of the return distribution. The fundamental variables, heuristic by their nature, come from the stock exchange, accounting, and other practices and have a distant relationship to the development of return distribution. All said above about the size and book-to-market ratio concerns these variables too.

All papers with fundamental variables use regression models to explain the excess returns by a set of factors that can be the company fundamentals directly, or the sensitives of each stock to systematic risk factors, usually denoted as “betas”. The latter case represents the settings of the classic asset-pricing models such as CAPM or the three-factor model of Fama and French (1993), hereafter FF3. Shapovalova et al. (2011) present a general regression model in the form:

$$r_{i,t} - r_t^f = c_t + \sum_{j=1}^{K} \gamma_t^j \theta_{i,t}^j + \omega_{i,t} \tag{9}$$

Here $r_{i,t}$ is the total return (the capital gain plus dividend yield) on stock i, $i \in (1,2,\dots,N)$, over period t, $t \in (1,2,\dots,T)$; r_t^f is the risk-free rate of returns over the period t, $\theta_{i,t}^j$ is the value of factor j, $j \in (1,2,\dots,K)$, of company i at time *t*, and γ_t^j is the return premium associated with a unitary increase in the value of factor *j* in period *t*. The term $\omega_{i,t}$ includes constant $\overline{\omega}$, time-effect term ω_t, and an error:

$$\omega_{i,t} = \overline{\omega} + \omega_t + \widetilde{\omega}_{i,t} \tag{10}$$

the error $\widetilde{\omega}_{i,t}$ is assumed to be a Gaussian zero-mean noise, *iid*. Shapovalova et al. (2011) make a step forward, taking account of the dynamic structure of the return premium in the form of the autoregressive random coefficient model:

$$\gamma_t^j = c^j + \alpha^j \gamma_{t-1}^j + \varepsilon_t^j \tag{11}$$

Here c^j and α^j are coefficients, ε_t^j is an error term. Choosing values for α^j, $|\alpha^j| \leq 1$, one can *post factum* adjust the dynamic structure of the model to the dynamic structure of the return premium observed in practice.

The regression approach to a process of unknown nature (e.g., the evolution of the return distribution, default risks, etc.) could hardly be efficient when researchers have to make a scientific guess about the explaining variables of the process under study. It is well known that regression cannot give any positive proof. It can only say that there are two or more concurrent processes, but it cannot reveal the cause-effect relationships between them or determine the space and time limits of those concurrencies. A vague vision of the asset return distribution and the excessive flexibility of proxies used in the models explain the mixed success of the results reported in literature and the absence of the best model. The researchers use the regression approach in their studies of asset pricing models because of a lack of more direct and efficient methods of attacking the problem. It does not mean that one must stop the regression studies, but support from other approaches to the problem could be helpful.

Now we return to reviewing papers from the first stream of literature. The firm size is the first heuristic factor introduced in CAPM by Banz (1981) to explain asset returns. Banz has noticed that smaller NYSE firms have, in general, higher returns for the same level of the beta than larger firms, and the more the difference in capitalization of the firms, the more the difference in returns. This relation suggests that the firm size (understood as a market value of the firm equal to the stock price times the number of outstanding stocks) captures some dimension of risk missed by the beta in CAPM. It makes the size factor a good candidate for an explaining variable in the asset pricing model. Another factor determining the cross-sectional variation of asset returns is the book-to-market ratio. Stattman (1980) and Rosenberg, Reid, and Lanstein (1985) are the first to report the book-to-market effect on the US stock market. Basu (1983) suggests the earning-to-price ratio of a stock as a new variable explaining asset returns, and Bhandari (1988) documents the effect of debt leverage on returns of common stocks. Fama and French (1992) claim that the size and book-to-market ratio have more predictive power on stock returns than the variables suggested by Basu and Bhandari. Fama and French (1993) include the size, book-to-market ratio, and beta as the determinants of the expected returns in their famous FF3. The papers (Fama & French, 1992; 1993) start the intensive search for a new, more precise CAPM. Fama and French (2004) put it as “the synthesis of evidence on empirical problems on the CAPM provided by (Fama & French, 1992) serves as a catalyst, marking the point when it is generally acknowledged that the CAPM has potentially fatal problems” (p. 36).

FF3 (Fama & French, 1993; 1996), following the above formalism, can be presented as

$$r_{i,t} - r_t^f = \gamma_t^M \beta_{i,t}^M + \gamma_t^{SMB} \beta_{i,t}^{SMB} + \gamma_t^{HML} \beta_{i,t}^{HML} + \omega_{i,t} \tag{12}$$

The factor *SMB* is the difference between the returns on diversified portfolios of small and big stocks. The factor *HML* is the difference between the returns on diversified portfolios of high and low book-to-market stocks. One can find the loadings $\beta_{i,t}^M$, $\beta_{i,t}^{SMB}$, and $\beta_{i,t}^{HML}$, by measuring the sensitivity of asset returns to the three factors from *N* time-series regressions for each stock (*N* is the number of stocks). The general opinion is that FF3 has more explanatory power than CAPM, especially in international and emerging markets (e.g., Drew & Veeraraghavan, 2002; Eraslan, 2013; Durga, 2020). However, this literature shows that the predictive power of *SMB* and *HML* on expected returns is unstable, varying significantly over time intervals and markets. FF3 has received much popularity among the economists, and many authors, who propose new methods for asset pricing, start with FF3. The discussion of FF3 continues till now (e.g., Allen & McAleer, 2018; Fama & French, 2018), revealing more and more difficulties in this model.

Daniel and Titman (1997) argue that firm characteristics rather than factor loadings on the *SMB* and *HML* portfolios determine the expected returns. They suggest an alternative stock characteristics model:

$$r_{i,t} - r_t^f = \gamma_t^M \beta_{i,t}^M + \gamma_t^{BtP} BtP_{i,t-l} + \gamma_t^{MCAP} MCAP_{i,t-l} + \omega_{i,t} \quad (13)$$

Here γ_t^{BtP} and γ_t^{MKAP} are the return premia generated by the firm's characteristics $BtP_{i,t-l}$ and $MKAP_{i,t-l}$ taken with a three-month lag denoted as l (BtP = Book-to-(Market)Price, MCAP (Market Capitalization = Size). Developing the method of firm characteristics, Shapovalova et al. (2011) explore the impact of the firm fundamentals on stock returns. Their list of accounting fundamental characteristics includes *PtB* = (Market) Price to Book ratio, *fPtE* = Projected Price to Earnings, *PtE* = Price to Earnings, *PtS* = Price to Sales, *PtCF* = Price to Cash Flow, *DY* = Dividend Yield, *fgEpS* = Projected Growth of Earnings per Share, *gEpS* = Growth of Earnings per Share, *gSpS* = Growth of Sales per Share, *fgSpS* = Projected Growth of Sales per Share, *IG* = Internal Growth, *MCAP* = Market Capitalization.

Shapovalova et al. (2011) select some fundamental variables used in the style analysis and business practice. These variables are the book-to-price ratio, market capitalization, sales growth, and internal growth. The authors confirm that style premia are time-varying, but their dynamic structure has occurred more complicated than the developed methodology can reveal. The authors claim that the reliability of their forecasts is on the level of results of Fama and French (1993; 1996), Daniel and Titman (1997), and other models using the unconditional approach.

Harvey et al. (2016) present a more detailed list of factors and models for cross-sectional anomalies in expected returns. The authors count hundreds of empirical models, distinguishing from one another with proxies for unobservable market factors. The authors have documented about 300 such proxies, and their number continues to grow. To find a better model in a huge set of available models, one must have an ordering parameter and a procedure for comparing competing empirical models. Shanken (1987) and Hansen and Jagannathan (1997) propose the measures of model misspecification. However, the intensive use of these measures has failed to separate the promising asset-pricing models. Gospodinov et al. (2013) argue that despite the econometric theory for comparing the empirical models based on the Hansen-Jagannathan distances (hereafter the HJ-distances), a general statistical procedure for model selection is still incomplete. Gospodinov et al. (2013) propose the methodology based on the pivotal Chi-squared versions of the model comparison tests that are easier to implement and analyze than their counterparts used in previous studies.

Nevertheless, it does not bring a decisive success. To explain this, Harvey et al. (2016) argue that using usual statistical significance cutoffs (e.g., *t*-ratio exceeding 2.0) in empirical tests is incorrect. Given the plethora of factors and the inevitable data mining, many historically discovered factors would be deemed “significant” by chance due to the rule of big numbers. To improve the situation, the authors propose three conventional multiple-testing frameworks and a new one that especially suits financial economics. While these frameworks differ in their assumptions, they are consistent in their conclusions. Harvey and his coauthors argue that a newly discovered factor to be accepted should have a *t*-ratio that exceeds 3.0. The main shortcoming of the analysis by Harvey et al. (2016) is their unconditional approach. Understanding its restrictions, the authors admit: “It is possible that a particular factor is very important in certain economic environments and not important in other economic environments. The unconditional test might conclude the factor is marginal" (Harvey et al., 2016, p. 7) When characteristics describing the asset depend on time, the unconditional models and tests can never be correct.

In our opinion, the failure of tests with HJ-distances used to select promising empirical models for asset pricing is due to the construction of the HJ-distances. They are derived using elliptic (symmetric) functions for the return distributions. (The main adventure of the elliptic distributions in this context is that they generalize the multivariate normal distribution and retain its analytical tractability irrespective of the number of assets.) The economic environment is assumed to be arbitrage-free. As a result, the return distributions remain symmetric, and the first and second moments determine the structure of both constrained and unconstrained HJ-distances (Gospodinov et al. 2012). Mark that the distribution returns used by Hansen and Jagannathan (1997) and Gospodinov et al. (2012; 2013; 2014) are the *market* return distributions having a more complicated structure than the return distribution of the firm.

Investors need to know the return distribution of each firm they plan to invest in: it gives them the mean returns on assets, variances and other statistical moments, and the default probability; this information helps investors optimize their portfolios. However, the investors do not know the return distribution of the firm; all they observe is the stochastic market returns. The investors can vaguely guess the return distribution of the firm by analyzing the time series of market outcomes. The market return consists of three parts: the dividends paid for a time unit, the change in the firm value for the same period, and the premium paid for the right of possessing that stock. The return distribution of the firm completely determines the first two terms. The eagerness of investors to have a particular stock in their portfolio affects the value of the last term. Two opposite tendencies govern this eagerness: the striving for wealth-maximizing on the one side and risk-avoiding on the other. These tendencies interfere with decision-making on the stock selection to the investor's portfolio, making the market return distribution different from the return distribution of the firm. The wealth-maximizing and risk-avoiding tendencies do not remain constant; one or the other behavior prevails at particular periods depending on the state of the stock and market. This subjective behavior of individual investors introduces additional stochastic shocks in the dynamics of market returns. The cumulative effect of all investors working with a particular stock at the stock exchange determines the market return distribution of this asset. Considering the structure and dynamics of market returns, it becomes clear that the market return distribution of any stock can hardly be symmetric.

Properties of the functions making the economic space of the firms paying BSEs are very different from that of the functions making the Hansen-Jagannathan space (the HJ space, Hansen & Jagannathan, 1997). Violation of the no-arbitrage principle in the HJ space happens due to misspecifications of empirical models. No-arbitraging in the space of firms paying BSEs can occur just as a rare random event. Therefore, the measures of these two spaces are different, and minimization of the HJ-distance between two empirical models with symmetric return distributions does not provide for the shortest distance between the models in the space of firms paying BSEs. This argument can explain the failure of attempts to select promising empirical models by minimizing the HJ-distance between the models.

CAPM and FF3 are static linear models with constant coefficients (betas). However, the economy permanently changes, experiencing boosts and recessions, and the firms have their business cycles, risk fluctuations, etc. Conditional asset-pricing models take account of this variability, making the expected returns depend on the information known to investors at a fixed moment. Jagannathan and Wang (1996) present the first conditional asset-pricing model. This model looks a bit naive compared to the sophisticated approach to the conditional models suggested by Hansen and Jagannathan (1997). Using 100 size-beta sorted portfolios of stocks from NYSE and AMEX between 1962 and 1990, Jagannathan and Wang claim that their conditional asset-pricing model can explain about 30% of the cross-sectional variations while static CAPM can explain about 1%. Another original contribution of this model to the asset-pricing theory is human capital used as a proxy for the market portfolio. The authors explain their choice by the high relative weight of the human capital in the economy.

The conditional approach to the asset-pricing modeling suggested by Hansen and Jagannathan (1997) and further developed by Gospodinov et al. (2012; 2013; 2014) presents a mature treatment of the asset pricing in time-varying conditions. See above the discussion on constructing this type of empirical asset-pricing models for the symmetric return distributions based on the HJ-distances.

Lettau and Ludvigson (2001) suggest a conditional consumption-based asset pricing model (a conditional consumption CAPM, or CCAPM). They start with recognizing the fact that expected returns are time-varying due to the changing economic environment. They look for the CCAPM as a linear function of fundamental variables with coefficients depending on time. In that model, the expected asset return is a function of its conditional covariance with a particular factor normalized by the conditional variance of this factor. To find the conditional variance and covariance, the authors model the dependence of parameters on time by scaling factors. Their scaling method is close to the method presented by Jagannathan and Wang (1996), who show that the performance of the asset-pricing model can be noticeably improved by conditioning the market factor on the default premium. Lettau and Ludvigson use another scaling factor, named *cay*, which is the difference between log consumption (*c*), the weighted average of log-asset wealth (*a*), and log-labor income (*y*). The authors argue that movements in *cay* are a good proxy for movements in the consumption-aggregate wealth ratio. They claim that the scaled multifactor version of conditional CCAPM performs well in explaining the cross-section of average returns: it explains nearly 70% of the cross-sectional variation in expected returns, about as well as FF3.

Conditional asset pricing models represent a new important step in the theory of asset pricing: the recognition of the effect of changes in the economic environment and their influence on asset pricing. Hansen and Jagannathan (1997) give a theoretical proof of the necessity to consider the time dependence of asset returns. Unfortunately, the first attempts to take account of the economic dynamics through scaling are not very successful: the models use not-so-good proxies for reflecting the time dependence of asset returns, and the heuristic method of scaling seems doubtful. The weaknesses of the conditional asset pricing models leave an open door for just criticism. However, the article (Lewellen & Nagel, 2006) rejects the very right of existence for the conditional models. Of course, discussing the weaknesses of conditional models, the authors refer to the existing models (Jagannathan & Wang, 1996; Lettau & Ludvigson, 2001; etc.), but they jump to the negative conclusion about *all* conditional models. The authors claim that they present an empirical proof that the conditional models cannot explain asset-pricing anomalies. They directly estimate the conditional coefficients using short-window regressions (a five-day window). "As long as betas are relatively stable within a month or quarter, then simple CAPM regressions estimated over a short window—using no conditioning variables—provide direct estimates of the conditional alphas and betas" (Lewellen & Nagel, 2006, p. 291). Rather than performing regression once, using the full-time series of returns, they estimate it separately every quarter using daily or weekly returns. The inevitable consequence of this approach is a significant increase in the number of regressions and corresponding errors. These errors drastically decrease the precision and accuracy of their pricing model. The authors try to verify low-quality conditional models with another low-quality model. That fruitless attempt makes the claim of Lewellen and Nigel on providing for an empirical proof that the conditional CAPMs do not explain the asset pricing anomalies unsubstantial and false.

Vassalou and Xing (2004) present another revision of FF3, analyzing the relationships between the size, book-to-market ratio, and the *default probability*. The authors try to estimate the default probability using the calibrated model of Moody's KMV (Bohn, 2006); they call their estimate the default likelihood indicator, *DLI*. Moody's KMV suffers dramatic drawbacks, producing very rough and misleading estimates of the firm default probability. First, to construct the key characteristic of Moody's KMV, the distance-to-default (*DD*), measuring the distance between the mean log-value of the firm and a specially selected default line, the authors start with the option pricing formula for the market value of the firm. Merton derives this formula to estimate the value of a zero-coupon bond; it does not fit the case of the equity value. Second, this formula is based on the assumption that the firm can default only at the maturity of its debt, while the firm can theoretically default at any time. Third, the authors assume that the firm makes no payments, and its log-value distribution remains normal, while this distribution is really skewed. Forth, it is assumed that *DD* uniquely determines the default risk. Using *DD* calculated for the normal distribution of the firm log-value and the default database, the model computes a share of firms with that *DD* that have defaulted within a year. This share has Moody's trademark name of Expected Default Frequency (*EDF*) and is a rough estimate of the default probability at a distance of one year. It is easy to understand that the assumption that *EDF* is a function of *DD* only is far from reality. Two firms with equal *DDs* at one moment can have different default probabilities because the firm distribution depends on the parameters of the firm and its business environment (the debt leverage, interest rate, inflation rate, taxation rate, etc.). *The listed drawbacks* are not specific for Moody's KMV; they are *immanent to all calibrated models using GBM*.

The drawbacks of Moody's KMV make it hardly advisable for estimating the default probability of an individual firm. However, one can apply this method to *qualitative analysis* of the cumulative market effects. To measure these effects, Vassalou and Xing introduce the aggregate default likelihood measure (*ADLM*), defined as a mean *DLI* over all firms in the market, and the aggregate survival rate, $ASR = 1 - ADLM$. The authors study the relations between the size, book-to-market ratio, and default risk. Constructing and testing a regression model with these variables, they conclude that the size effect exists only within the market segment that contains the stocks with the highest default risk. The authors suppose that the high mean returns of small stocks compared to the rest of the market are the compensation for the high default risk they take. Considering the book-to-market effect, Vassalou and Xing document that this effect is significant only in the quintiles with high default risks. They also register a monotonic relation between default risk and the book-to-market factor in these quintiles. This relation is not monotonic in other quintiles with lower default risks. (See our commentary on the size and book-to-market factors in Section 4). Discussing the effect of default probability, the authors believe that the default risk is priced and must be included in asset pricing models. They explain the higher returns of small firms as a compensation for their higher default risk compared to the risk of big firms. Section 4 shows that the higher the default probability, the lesser the returns, the greater the losses. The default risk cannot be a priced variable making a benign effect on asset returns. Another objection against the default risk as an explaining variable in the asset-pricing model is the inability of the authors to compute a reliable value of the default risk in their model.

From the discussion above, it is clear that attempts to construct a new CAPM by guessing the variables that improve a predictive power of the single-factor CAPM are not productive. A more systematic approach using information about the return distribution of the firm could be more helpful. Second, wrong ideas about the symmetry of return distribution impede the development of efficient asset pricing models. Knowledge of the skewed return distribution provided by EMM can lead the scientific search to more perspective methods of solving the problem. We show in Section 4 that the variance is, in general, a poor measure of the default risk for skewed return distributions. It means that the perspective asset pricing model with higher moments will not include Sharpe's beta as an explaining variable.

Now we consider the stream of the mean-variance-skewness investigations. Harvey and Siddique (2000) examine the linkage between the factors in asset-pricing models and systematic co-skewness. In the single-factor model, investors estimate the security by its mean return and variance. However, the return distribution in the market demonstrates some skewness. Everything else being equal, investors prefer right-skewed portfolios to left-skewed ones because the right-skewed portfolios promise them higher returns (Kraus & Litzenberger, 1976). The authors assume that portfolio skewness is a consequence of the limited liability in all equity investments that may induce option-like asymmetry in returns. Another cause of portfolio skewness suggested by Harvey and Siddique (2000) is that the managers prefer portfolios with high positive skewness. The authors test a three-moment conditional asset-pricing model to see if the stocks with big negative co-skewness with the market will earn higher risk premia. (See our comment on the sign of co-skewness in the discussion of results by Barone-Adesi et al. (2004).) They use the size and book-to-market factors as proxies for the co-skewness factor because direct co-skewness estimating is difficult. Harvey and Siddique (2000) document that skewness varies among portfolios of different sizes and book-to-market levels. The authors find skewness very helpful in explaining the cross-sectional variations of equity returns. They also show that the momentum effect (Jegadeesh & Titman, 1993) is related to systematic portfolio skewness.

Longin and Solnik (2001) study the conditional correlation structure of international equity returns using a statistical method based on the extreme value theory. Their method tests whether these correlations deviate from what one would expect under multivariate normality. They show that high volatility with high absolute returns does not lead to an increase in conditional correlation. It is only during market downturns (the bear markets) that conditional correlation strongly increases; conditional correlation does not seem to grow during market upturns (the bull markets). This result makes another empirical proof that the fear of losing wealth is a much stronger emotion for investors than their striving to increase wealth.

Patton (2004) considers the effects of two types of asymmetries in the joint distribution of stock returns on asset allocation and portfolio optimization. The first type is skewness in the distribution of individual stock returns (e.g., Kraus & Litzenberger, 1976; Friend & Westerfield, 1980; Singleton & Winberger, 1986; Richardson & Smith, 1993; Harvey & Siddique, 2000). The second asymmetry concerns a higher correlation of stock returns during market downturns than during market upturns (Erb et al.,1994; Longin & Solnik, 2001; Ang & Chen, 2002; Bae et al., 2003). The author examines the problem of an investor with a constant relative risk aversion (CRRA) allocating wealth between the risk-free asset, the small- and large-cap indices comprised of the first and tenth deciles of the US stocks sorted by market capitalization. Using the copula technique and CRRA utility function, Patton (2004) demonstrates that skewness in individual stock returns and the asymmetric correlation of stock returns in market downturns induce the negative skewness in the distribution of portfolio returns. (EMM modeling and empirical observations show that portfolio returns can have both negative and positive skewness.) Patton documents evidence that the model capturing skewness and asymmetric correlation dependence yields better portfolio decisions than the bivariate normal model. The difference in returns of the most flexible density model and the bivariate normal model makes between 0 and 27 basis points per year.

Barone-Adesi et al. (2004) consider market co-skewness and investigate its role in testing the asset pricing models. They start with constructing a quadratic model of the mean excessive returns; the model includes the market returns and the square of market returns as the two factors. Following Harvey and Siddique (2000), they define an asset as having positive co-skewness with the market when the residuals of regression on a constant and the market return positively correlate with the squared market return. With this definition, an asset with positive (negative) co-skewness reduces (increases) the portfolio risk to high market returns and commands lower (higher) expected returns in equilibrium. They find that small firms have negative co-skewness while large firms have positive co-skewness. It is worth noting that this definition of the positive/negative co-skewness is opposite to the generally accepted definition of skewness as the third statistical moment of the asset return distribution. Of course, co-skewness and skewness of returns have different numeric values, but they must have the same sign. According to EMM, small and large firms may have both negative and positive skewness in their returns, but small firms have more often positively skewed returns, while large firms have mostly negatively skewed returns.

Besides evaluating asset pricing models with co-skewness, Barone-Adesi et al. (2004) investigate the consequences of asset pricing tests without co-skewness. The authors empirically prove that some portfolio characteristics, such as the firm size, explain expected excess returns (e.g., Banz, 1981; Fama & French, 1992). If the market co-skewness correlates with the size, it has spurious explanatory power for the cross-section of expected returns because the size factor proxies for omitted co-skewness. The authors empirically prove that co-skewness and the size factor are correlated. This finding suggests that the empirically observed relationship between the size and the returns may be explained by such a systematic risk factor as market co-skewness; see also Harvey and Siddique (2000, p. 1281).

Jondeau and Rokinger (2006) address the problem of how the non-normality of asset returns impacts the allocation of wealth for utility-maximizing investors. Using a Taylor series expansion, they approximate the utility as a function of three or four moments and compute the optimal portfolio allocation. For optimization, the authors use a distance measure between portfolio weights and estimate the opportunity cost of a sub-optimal strategy. Jondeau and Rokinger (2006) document that when the returns show slight deviation from normality, different allocation strategies provide almost the same allocation, suggesting that the mean-variance method correctly approximates the expected utility. When the return distribution severely deviates from normality, the extension to the three- or four-moment criterion results in much better approximation of the expected utility than the mean-variance criterion. The goodness of approximation is highlighted by the opportunity cost which is very small in all instances for the four-moment criterion. In conclusion, Jondeau and Rokinger (2006) remark that for further development of conditional asset allocation, one must have a model for returns “with an asymmetry and fat tails”. EMM is the very model.

Portfolio Optimization

Portfolio selection with higher moments has received considerable attention and provided many good results in the last forty years. The seminal paper (Kraus & Litzenberger, 1976) marks the beginning of systematic work on portfolio allocation and asset pricing with higher moments. The authors develop the asset pricing model in which the expected asset return is a linear function of the Sharp beta and co-skewness of the asset with the market portfolio. This result follows from the assumption that the utility of portfolio return is a function of the portfolio expected value, variance, and skewness. Following the logic of Kraus and Litzenberger (1976), portfolio allocation as a result of the tradeoff between risk aversion and maximization of the expected return by choosing higher skewness requires the specification of a suitable utility function $U(X_p)$ and computation of its expected value. To perform this, one must determine the distribution of the portfolio return. Without it, the researchers look for a suitable multivariate distribution of asset returns for deriving the required distribution for the portfolio return.

Athayde and Flores (2004) are the first who consider the problem of choosing an efficient portfolio by taking account of three moments of the asset return distribution: the mean, variance, and skewness. Setting aside the problem of determining the asset moments and assuming that the moments are known precisely, they study a portfolio consisting of *n* risky assets plus one riskless asset and derive a formula for the solution surface. They document the geometric properties of the efficient surface in the space of three moments. The authors obtain a shape of the mean-variance-skewness efficient frontier for some examples using a simulation technique. Athayde and Flores (2004) show that the mean-variance solution of Markowitz can be most inefficient when the asset skewness is significant.

The authors, sharing common vision of the skewness role, believe that the optimal solution includes the unconstrained maximization of portfolio skewness. However, high positive skewness comes together with a high default probability (see Section 4); it makes the "optimal" portfolio profitable but highly unstable. To improve the situation, one must look for the maximum of the mean portfolio return and skewness under the constraint that the default probability of each asset to be included in the portfolio remains below an agreed threshold DP_{Tr}. After that, one must continuously check and recheck the mean returns and default probabilities of assets keeping the portfolio returns and risks under control. Of course, the authors cannot find the default probability of each asset in their problem settings, and EMM could be of great assistance in this case.

It is worthy to emphasize that in the higher-moment problem, the variance is not a variable uniquely related to the default probability of an asset as it is in the mean-variance portfolio optimization, when the return distribution follows GBM and is symmetric. Therefore, one must change the mean-variance optimization of Markowitz for another problem of maximizing the mean return and skewness, given the default probability (or the intensity of default probability) is no more than an agreed value. The default probability is a crucial characteristic of any asset or portfolio irreducible to any statistical moment or a combination of statistical moments.

Krueger (2021) considers the problem of portfolio optimization using a nonparametric method of the data envelopment analysis (e.g., Morey & Morey, 1999; Edirisinghe & Zhang, 2007) in the interpretation of Briec and Kerstens who suggest a mean-variance-skewness analysis of portfolio optimization (e.g., Briec et al., 2004; Briec et al., 2007; Briec & Kerstens, 2010). Following the ideas of Jondeau and Rokinger (2006), the author relates higher moments of the return distribution with the utility of the investor holding a particular portfolio of assets. Realizing portfolio optimization, he starts with a nonparametric calculation of statistical moments by averaging asset returns over time:

$$\widehat{m}_i = T^{-1}\sum_{t=1}^{T} R_{it} \;,\; \hat{v}_{ij} = T^{-1}\sum_{t=1}^{T} r_{it}r_{jt} \;,\; \hat{s}_{i,n(j-1)+h} = T^{-1}\sum_{t=1}^{T} r_{it}r_{jt}r_{ht} \tag{14}$$

$$\hat{k}_{i,\,n^2(j-1)+n(h-1)+l} = T^{-1}\sum_{t=1}^{T} r_{it}r_{jt}r_{ht}r_{lt}\ ,\ r_{it} = R_{it} - \hat{m}_i, i = 1,2,\dots,n$$

R_{it} is the return on *i* asset at time *t*, $\hat{m}_i$ is the estimated mean return on asset *i*, $\hat{v}_{ij}$ is estimated covariance between assets *i* and *j*, $\hat{s}_{i,n(j-1)+h}$ is co-skewness for assets *i*, *j*, and *h*, and $\hat{k}_{i,\,n^2(j-1)+n(h-1)+l}$ is co-kurtosis for assets *i*, *j*, *h*, and *l*; *n* is the number of stocks. Corresponding portfolio characteristics are: ***m***—the vector of portfolio returns, ***V***—the portfolio covariance matrix, ***S***—the portfolio skewness matrix, and ***K***—the portfolio kurtosis matrix. One can see from the notation of moments that the author supposes the vector of returns and all matrices to be time independent. To secure the necessary precision and accuracy of portfolio optimization, time *T* must be chosen sufficiently large. It makes the matrices ***V***, ***S***, and ***K*** huge. To deal with them, the author applies the method of matrix shrinkage developed by Ledoit and Wolf (2003; 2004), Martellini and Ziemann (2010), Boudt, Cornilly, and Verdonck. (2020).

For an investor holding a portfolio of assets with shares $0 \le x_i \le 1$, it makes the portfolio vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n)'$, $\boldsymbol{1}'\boldsymbol{x} = 1$, $\boldsymbol{1}' = (1,1,\dots,1)_n$, the portfolio return $R(\boldsymbol{x}) = \boldsymbol{x}'\hat{\boldsymbol{m}}$ with the expected return $m_p(\boldsymbol{x}) = E\big(R(\boldsymbol{x})\big) = \boldsymbol{x}'\boldsymbol{\mu}$, the portfolio covariance matrix $v_p(\boldsymbol{x}) = Var\big(R(\boldsymbol{x})\big) = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$, the skewness matrix $S_p(\boldsymbol{x}) = Skew\big(R(\boldsymbol{x})\big) = \boldsymbol{x}'\boldsymbol{S}(\boldsymbol{x}\otimes\boldsymbol{x})$ (symbol $\otimes$ denotes the Kronecker product), and the kurtosis matrix $K_p(\boldsymbol{x}) = Kurt\big(R(\boldsymbol{x})\big) = \boldsymbol{x}'\boldsymbol{K}(\boldsymbol{x}\otimes\boldsymbol{x}\otimes\boldsymbol{x})$.

Then the author introduces the representative set *D* and the admissible set *P*:

$$D = \{z \in R^{|z|}: \boldsymbol{x} \in P, z \le z_p(\boldsymbol{x})\}\ ,\ P = \{\boldsymbol{x} \in R^n: \boldsymbol{x} \ge \boldsymbol{0}, \boldsymbol{1}'\boldsymbol{x} = 1\} \tag{15}$$

|*z*| denotes the dimension of vector *z*: |*z*| = 2 for MV problem, |*z*| = 3 for MVS problem, and |*z*| = 4 for MVSK problem.

The shortage function or the directional distance function which serves as a measure of portfolio efficiency is defined as

$$S_g(\boldsymbol{x}) = \sup\{\delta \ge 0: z_p(\boldsymbol{x}) + \delta\boldsymbol{g} \in D\}. \tag{16}$$

Vector ***g*** determines the direction in which the moments should be expanded or contracted: in MV case $g_1 \ge 0$ and $g_2 \le 0$ for expending expected returns and contracting their variance. To find a point on the portfolio frontier along the direction ***g***, one must maximize the measure of portfolio efficiency together with ***x***. For a point on the frontier $\delta = 0$.

The author put the portfolio optimization problem, using the directional distance function, as (Krueger, 2021, p. 1441):

$$\begin{gathered}
\max_{\delta,\boldsymbol{x}}\ \delta \\
\text{s.t. } \hat{m}_i + \delta g_m \le \hat{m}_p(\boldsymbol{x}) \\
\hat{v}_i - \delta g_v \ge \hat{v}_p(\boldsymbol{x}) \\
\hat{s}_i + \delta g_s \le \hat{s}_p(\boldsymbol{x}) \\
\hat{k}_i - \delta g_k \ge \hat{k}_p(\boldsymbol{x}) \\
\boldsymbol{1}'\boldsymbol{x} = 1,\ \boldsymbol{x} \ge \boldsymbol{0},\ \boldsymbol{g} = (g_m, g_v, g_s, g_k)',\ i \in \{1,2,\dots,n\}
\end{gathered} \tag{17}$$

He solves this problem for the series of single fixed directions first by maximizing the mean return (choosing vector $\boldsymbol{g} = (|\hat{m}_i|,0,0,0)'$), then by minimizing only the variance ($\boldsymbol{g} = (0,\hat{v}_i,0,0)'$), then maximizing the skewness ($\boldsymbol{g} = (0,0,|\hat{s}_i|,0)'$), and minimizing kurtosis ($\boldsymbol{g} = (0,0,0,|\hat{k}_i|)'$) (Krueger, 2021, pp. 1442, 1450-1452).

For the empirical application of the developed technique, Krueger chooses the five-factor model (Fama & French, 2015). For the market returns, he uses returns on indices of US firms listed on the NYSE, AMEX, or

NASDAQ. Then he develops portfolios formed on the size as measured by equity quintiles, including all NYSE, AMEX, or NASDAQ stocks, portfolios formed on quintiles of book-to-market ratio, portfolios formed on quintiles of operating profitability (annual revenues less cost of goods sold, commercial, interest, general, and administrative expenses divided by book equity), and portfolios formed on quintiles of investments (the change in total assets from the fiscal year (t - 2) to the fiscal year (t - 1), divided by the total asset in the year (t - 2). The author documents that the efficient portfolios are composed of small stocks, stocks with a high book-to-market ratio, stocks with high operating profitability, and stocks of firms with low investments. Krueger confirms the dramatic role of skewness in portfolio optimization; skewness helps speed up the search for efficient frontiers.

With due respect to the sophisticated model developed by the author and his predecessors, we call the reader to take the conclusions of this paper with great caution. First, the author supposes the statistical moments of a firm to be independent of time (Eq. (13)). However, the moments *are* the functions of time (see Section 4), and the time-averaging procedure introduces unacceptable errors in the moment estimations. Huge volumes of measurements that one needs for estimating higher moments (the variance, skewness, kurtosis) in the nonparametric study make this method completely unreliable. The second weakness of this paper is the unconditional maximization of skewness when solving Eq. (17). Positive skewness comes with a high default probability (see Section 4) that makes the “optimal” portfolio profitable but unstable. Mark that all portfolios that the author takes for the efficient ones (small size, large book-to-market ratio, low investments) are of low stability. To remove this defect, one must reset the maximization problem like this: find the maximum mean return and skewness under the constraint that the default probability does not exceed an agreed threshold.

As we have said above, for the skewed return distributions, the variance is not an adequate measure of default probability; thus, the minimization of variance becomes useless. Unfortunately, this method, as it is, cannot estimate the default probability of the firm. EMM can be of great assistance in solving that task. However, even this help cannot save the method because of its first deadly weakness. In our opinion, the nonparametric methods of solving the optimal portfolio allocation problem have a very little chance for success. The skew-elliptic methods (see below) seem much more promising.

Another stream of literature on optimal portfolio selection includes papers that introduce asymmetry in the standard GBM distribution of asset returns, considering the return dynamics as a jump-diffusion process. Since the original paper (Merton, 1971), the problem of portfolio choice in the presence of asset skewness has become a topic of growing interest (e.g., Merton, 1976; Aase, 1984; Aase & Oksendal, 1988; Liu et al., 2003; Das & Uppal, 2004; Cvitanic et al., 2008). We discuss here a pair of typical papers from this block.

The paper (Liu et al., 2003) studies the implications of event-related jumps in security prices and volatility on optimal dynamic portfolio strategies. In the model, both the security price and the volatility of its returns follow jump-diffusion processes. Jumps are triggered by Poisson events, which have the intensity proportional to the level of volatility. The authors obtain a closed-form solution for the optimal portfolio weight in these conditions. They show that the optimal portfolio strategy with event-generated jumps makes an investor behave if she faces short selling and borrowing constraints even if there is none. The authors conclude that the optimal portfolio is a blend of the continuous-time GBM solution and the optimal portfolio for the static optimal buy-and-hold problem. To illustrate theoretical findings, the authors provide two examples. The first is the model where the risky asset follows a jump-diffusion process with deterministic jump sizes and constant return volatility. The model demonstrates that the investor facing jumps may choose a portfolio very different from the portfolio optimal for the pure diffusion process of the return development. In the presence of jumps, the investor prefers to see less risky assets in her portfolio. The second model considers the case when the risky asset and its return volatility follow a jump-diffusion process with deterministic jump sizes. The authors conclude that the optimal portfolio weight does not depend on the level of volatility but mainly depends on the horizon of investments. They conclude that volatility jumps significantly affect the optimal portfolio above and beyond the effect of price jumps.

Das and Uppal (2004) aim to evaluate the gains from international diversification in the presence of systemic risk, understood as the risk from infrequent events that strongly correlate across a large number of assets. Such connected shocks across international stock markets result mainly from information transmission (Engle et al., 1990). Das and Uppal (2004) construct a dynamic model with preferences given by a constant relative risk aversion utility function (CRRA), where the effects of the return skewness appear due to a standard jump-diffusion process describing the asset return development. The authors assume that the jumps across assets happen synchronously, though their size can vary across assets. They find that while the systemic risk causes a shift in the allocation of wealth from risky assets to riskless assets, it has a negligibly small effect (it decreases the $1 return by $0.001 for the developed-country indexes and $0.06 for the emerging-country indexes). In these problem settings, the portfolio taking account of the systemic risk just insignificantly differs from the portfolio using the mean-variance optimization of Markowitz. Das and Uppal forget that portfolios and their assets are skewed even without systemic jumps, and the mean-variance optimization is not generally efficient.

As one can see, in the jump-diffusion approach to the optimal portfolio allocation, the skewness in the asset returns is provided by the jumps generated and controlled by the market events exterior to the firm dynamics. Depending on the intensity and size of jumps, their effect on portfolio allocation varies significantly. However, the asset returns have their own asymmetry generated by the firm dynamics additionally strengthened by a collective behavior of market traders. Because the return distribution is governed by Eq. (1), and market events are exogenous to the firm development, one can never correctly describe the asymmetry of the naturally skewed return distribution (Eqs. (1) and (6)) by the asymmetry generated by the jump-diffusion process. Thus, one must take cautiously all conclusions of the papers using jump-diffusion processes.

Another branch of the studies in optimal portfolio allocation contains papers using novel probability distributions with controllable skewness, which can flexibly describe both empirical data and theoretical results in asset returns. These distributions belong to the class of analytically tractable skew-elliptical probability distributions. The first such distribution is the multivariate skew-normal distribution (MSN) developed by Azzalini and Dalla Valle (1996). Its further development, having more controlling parameters and, thus, more flexible, has appeared in the papers (Arnold & Beaver, 2000; Adcock & Shutes, 2001) under the name of the multivariate extended skew-normal distribution (MESN). A more powerful generalization of these two distributions (but much more complicated) is the multivariate skew-Student *t* distribution having heavier tails (Branco & Dey, 2001; Adcock, 2002; Azzalini & Capitanio, 2003). One can find comprehensive reviews of these probability distributions in the papers (Azzalini, 2005; Adcock et al., 2012).

One method of generating the skew-normal distribution is to take a bivariate normal distribution (X, Y) and then develop a conditional distribution X given the condition $Y > 0$. This representation is equivalent to the distribution of $U + \lambda V$, where U is a normal distribution and V, independent of U, is the normal distribution $N(\tau, 1)$ truncated from below at zero. This distribution is analytically tractable and possesses a lot of useful theoretical properties. Another way to introduce a skewed distribution is, for example, to consider a random U having normal distribution and a non-negative V having exponential distribution, producing the normal-exponential distribution (Aigner et al., 1977). Without information about the behavior of asset return distribution, the researchers studying the optimal portfolio allocation have to make reasonable assumptions about this behavior. The goodness of such distributions depends on how correct the assumptions are.

Harvey et al. (2010) address the problem of the optimal portfolio allocation using higher moments (skewness) and estimating risks in a coherent Bayesian framework. In a comprehensive study, they specify a Bayesian probability model for the joint distribution of asset returns. In the next step, the authors calculate and maximize expected utilities in the face of parameter uncertainty using Bayesian methods. Discussing the role of skewness, the authors share a general opinion of the unconditionally beneficial effect that positively skewed returns make on portfolio selection. For the probability distribution describing asset returns, the authors choose the multivariate skew normal probability distribution suggested by Sahu et al. (2003) and then offer a generalization of this model.

The multivariate skew normal distribution is presented by a mixture of an unrestricted multivariate normal density and a truncated, latent multivariate normal density $X = \mu + \Delta Z + \varepsilon$, where μ and Δ are unknown parameter vector and matrix, ε is a normally distributed error vector with a zero mean and covariance Σ, and Z is a vector of latent random variables. Z comes from the multivariate normal with the mean zero and an identity covariance matrix and is truncated at zero from below ($Z \geq 0$). Harvey et al. (2010) consider Δ as an unrestricted random matrix allowing for co-skewness. Under reasonable assumptions about the properties of μ, Δ, and Σ, the estimation is done using the Markov Chain Monte Carlo algorithm. The authors use Bayesian methods to include parameter uncertainty into the predictive distribution of returns and then optimize the expected utility. The optimal portfolio allocation is based on the maximization of asset mean returns and skewness keeping the variance fixed according to the assumption that investors prefer assets with a high upside to assets with a large downside (Kraus & Litzenberger, 1976). As we know, this allocation leads to the development of the profitable but unstable portfolio threatening its investor with sudden losses.

The important paper (Mencia & Sentana, 2009) makes the mean-variance-skewness analysis fully operational by working with another family of multivariate asymmetric distributions known as Location-Scale Mixtures of Normals (LSMN). This family nests several important symmetric distributions as the Gaussian and Student *t* distributions and asymmetric distributions like the Generalized Hyperbolic (Barndorff-Nielsen, 1977), Normal-Gamma (Madan & Milne, 1991), etc. LSMN nests the finite mixture of normals, which make a flexible and empirically plausible instrument to deal with asymmetric asset returns while remaining analytically tractable. The authors consider the portfolio allocation problem when asymmetric asset returns are assumed to be precisely known (no stochastic noise). An LSMN random vector $\boldsymbol{r}$ can be presented as

$$\boldsymbol{r} = \boldsymbol{v} + \xi^{-1}\boldsymbol{Y}\boldsymbol{\delta} + \xi^{-1/2}\boldsymbol{Y}^{1/2}\boldsymbol{\varepsilon} \ , \tag{18}$$

where $\boldsymbol{r}$, $\boldsymbol{v}$, and $\boldsymbol{\delta}$ are N-dimensional vectors, $\boldsymbol{Y}$ is a positive definite matrix of order N, $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \boldsymbol{I})$, and ξ is an independent positive mixing variable whose distribution depends on a vector of q shape parameters $\boldsymbol{\varrho}$. Since $\boldsymbol{r}$ given ξ is Gaussian with the conditional mean $\boldsymbol{v} + \xi^{-1}\boldsymbol{Y}\boldsymbol{\delta}$ and covariance matrix $\xi^{-1}\boldsymbol{Y}$, then $\boldsymbol{v}$ and $\boldsymbol{Y}$ act as the location vector and the dispersion matrix. Parameters $\boldsymbol{\varrho}$ allow for flexible tail modeling, while the vector $\boldsymbol{\delta}$ introduces skewness in this distribution. The authors prove that if the asset return distribution can be presented as an LSMN distribution, then the distribution of any portfolio combining those assets can uniquely be represented by its mean, variance, and skewness. From the investor's point of view, such portfolios have a very attractive property: any two portfolios can be expressed in terms of the first three moments because all higher-order moments fully depend on these three and tail parameters $\boldsymbol{\varrho}$. Mencia and Sentana (2009) show that the efficient part of the portfolio frontier can be spanned by three funds: the fund that together with a fund of riskless assets generates the usual mean-variance efficient frontier plus a fund that can be interpreted as the skewness-variance efficient portfolio. This result causes our strong doubts.

One can see from the solution of the optimal portfolio problem that the authors interpret the variance as an adequate measure of asset risks. We repeat that for skewed distributions, the default probability is an adequate measure of risk rather than the variance (compare Figures 9 and 12). The portfolio optimization problem with higher moments must be put like this. Determine the vector of asset weights providing for the maximum of the portfolio mean return and skewness given that the intensity of the firm default probability is under a given threshold. There is no efficient mean-variance portfolio and no efficient skewness-variance portfolio in this optimization problem. Unfortunately, the method used by the authors does not allow computing the intensity of default probability and the default probability. EMM could be of assistance in this case.

Modeling Return Distribution with EMM

We solve EMM problem (1.1)-(7.1) with $\pi(t) \equiv 1$ and $DL = 0$ (the unlevered firm) numerically, estimating the mean log returns $H\_BP(t)$, mean returns $MX\_BP(t)$, variance $VAR\_BP(t)$, skewness $SK\_BP(t)$, and default probability $DP(t)$. We trace down the dependence of statistical moments of the return distribution and the default probability on the parameters of the firm and its business environment. At this stage, we suppose that all *perfect market assumptions hold*. Model parameters are $R = 0.10$, $\sigma_0^2 = 0.02$, $C^2 = 0.01$, $T = 25$ (years).

The log return $x = \ln(RX/P_0)$, the (dollar) return $X = (P_0/R)e^x$, the log-return mean $H\_BP(t) = \int_{DL}^{\infty} x\hat{V}(x,t)dx$, the mean non-dimensional return $MX\_BP(t) = P_0/(R\langle X_0\rangle)\int_{DL}^{\infty} e^x\hat{V}(x,t)dx$. The initial distance to the default line is $H_0 = \int_{DL}^{\infty} xV_0(x)dx$, $V_0(x) = N(x; H_0, \sigma_0^2)$, the normal density. As one can see from the pictures, the return distribution and its derivatives depend on the parameter H_0; to a much lesser degree, the distribution depends on the initial variance σ_0^2 and the intensity of shocks C^2. The last two dependencies are not shown.

Returning to the question of the effect of firm size $\langle X\rangle$ on its mean returns, one can see that it is not the firm size that governs the firm state and dynamics, but the relationship between the mean yearly income $R\langle X(t)\rangle$ and yearly expenses $P(t)$. The same size $\langle X_0\rangle$ can lead the firm to a steady state with $H_0 > 1.5$ as well as to a shaky state with $H_0 < 1.3$. Therefore, a very popular parameter *SMB* (the small firm minus the big firm) does not uniquely determine the states of two firms and can refer to the difference in characteristics of two shaky firms, two steady firms, or steady less shaky and vice versa. This confusion explains the mixed success of the predictive power of size in asset pricing models. The same could be said about all proxies used in asset pricing models with "fundamental" variables. We will use the term "the small firm" referring to the shaky firm with a short distance to the default line, and "the big firm" meaning a steady firm with a long distance to the default line.

Figures 1 and 2 show that the small firm has a lesser return than the large firm given that they have the same expected rate of returns ($\langle X_S\rangle/P_S < \langle X_B\rangle/P_B$). It is clear that to achieve a greater return than a large firm the small firm must take a project with a higher rate of expected returns, R, and a higher level of risk, C^2. Of course, this tactic leads to a higher death rate for small firms. However, the conditional distribution, consisting of the firms who survived at the time of observation *t*, will show higher returns for the small firm compared to the large one. Large firms usually value the stability of their development rather than the fast growth of their assets and, therefore, they choose projects with a lesser rate of expected returns.

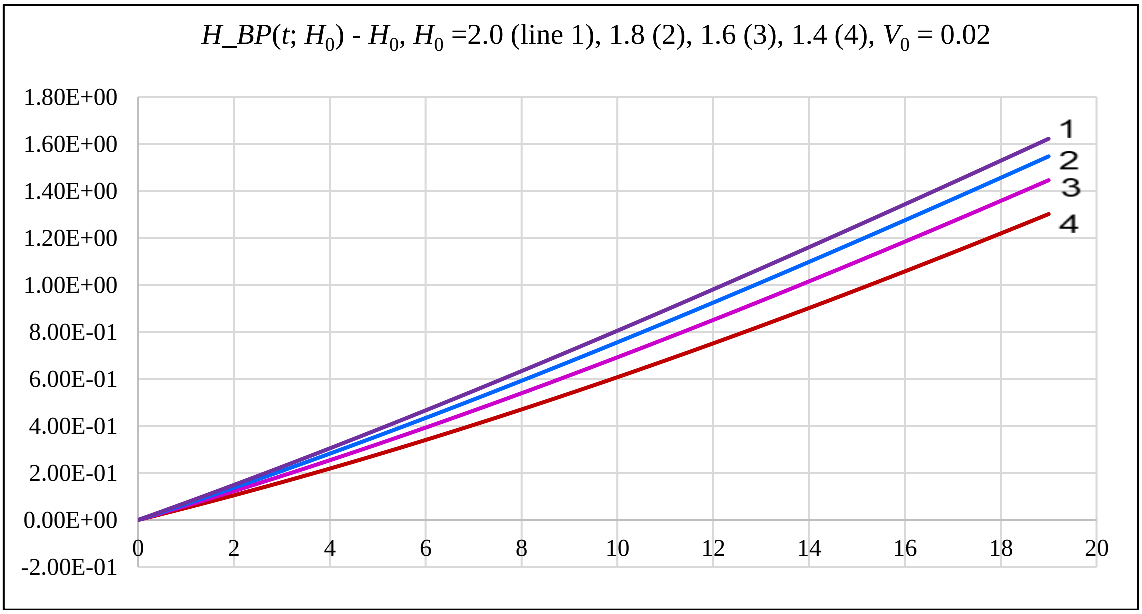

Figure 1. Mean log returns *H_BP*(t) - H_0 as a function of time (years) for various initial conditions H_0 = 2.0 (line 1), 1.8 (line 2), 1.6 (line 3), and 1.4 (line 4); VAR_0 = 0.02.

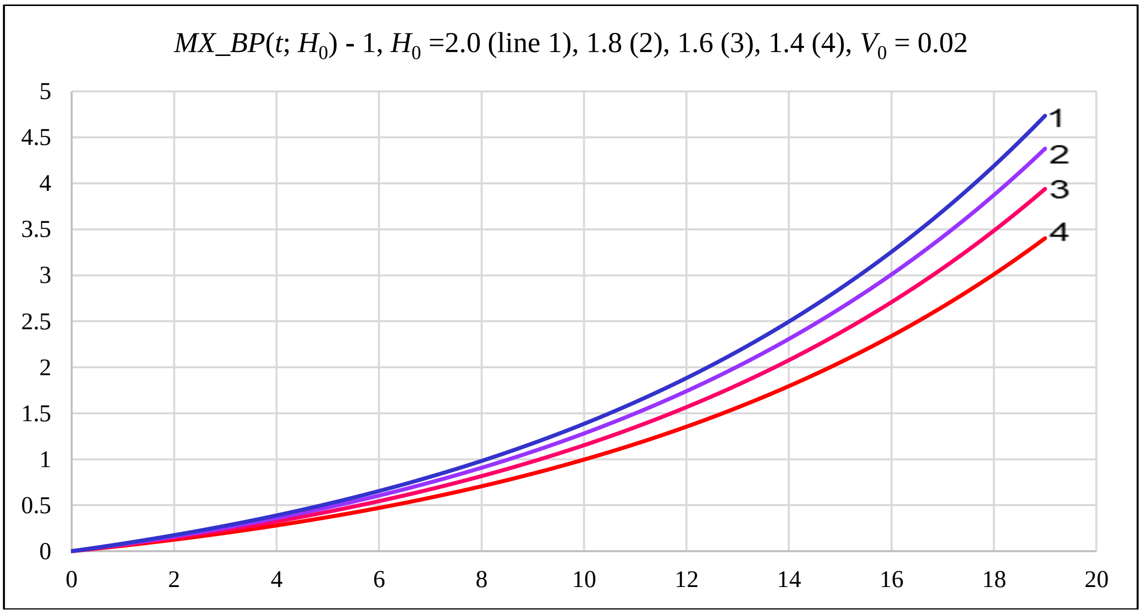

Figure 2. Mean non-dimensional returns *MX_BP*(t) - 1 as a function of time (years) for various initial conditions H_0 = 2.0 (line 1), 1.8 (line 2), 1.6 (line 3), and 1.4 (line 4); VAR_0 = 0.02.

Figure 3 demonstrates the time dependence of the variance *VAR_BP*(t) - VAR_0 for different H_0 values of steady (large) firms. Mark how the variance grows with the decrease of the distance to default because of faster growth of the distribution skewness. The variance and the default probability are concurrent, and the variance can be considered as a measure of risk. Figure 9 shows the time dependence of the variance *VAR_BP*(t) - VAR_0

for shaky (small) firms. There is no clear relationship between the initial state of the firm and the variance in this case, while the default probability grows as the distance to default decrease. The inference from this fact is that the variance is no measure of risks for shaky firms. It means that the variance (beta) cannot be an explaining variable in a perspective asset-pricing model.

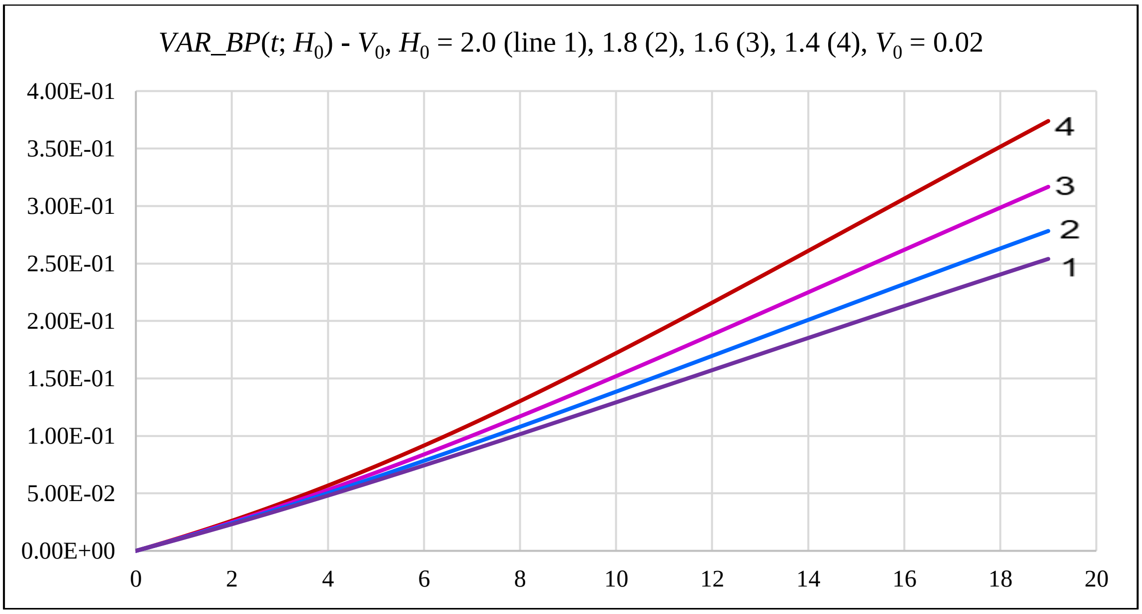

Figure 3. Variance $VAR\_BP(t)$ - VAR_0 as a function of time (years) for various initial conditions H_0 = 2.0 (line 1), 1.8 (line 2), 1.6 (line 3), and 1.4 (line 4); VAR_0 = 0.02.

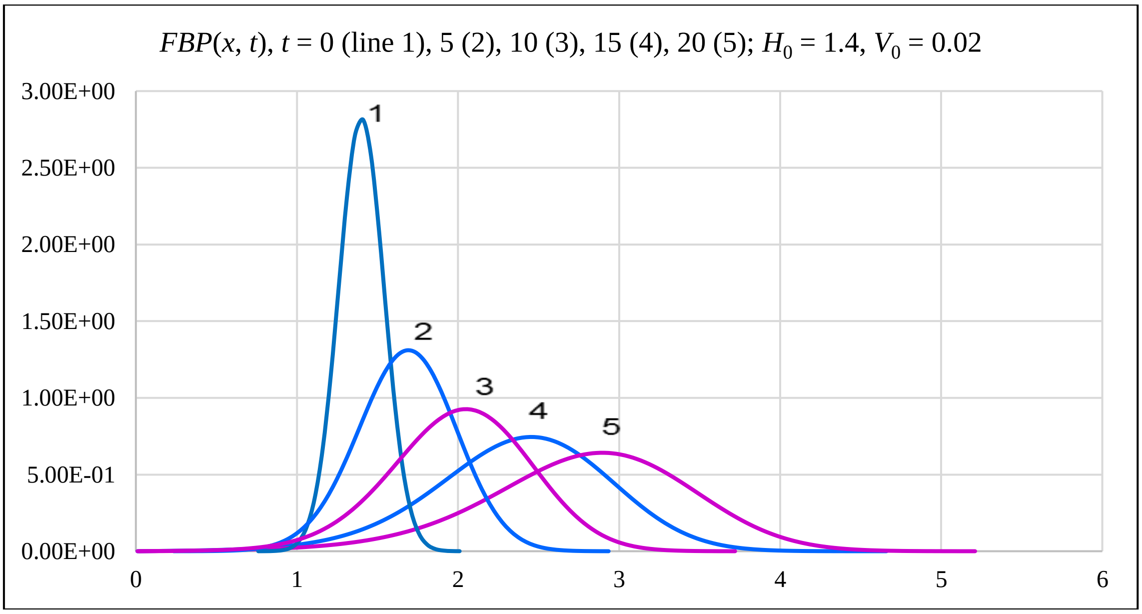

Figure 4. Evolution of the boundary problem distribution $FBP(x, t)$ for H_0 = 1.4 and t = 0 (line 1), 5 (line 2), 10 (line 3), 15 (line 4), 20 (line 5); VAR_0 = 0.02.

Figures 5, 6 and 11, 12 present the development of distribution skewness and the default probability for different initial distances to the default line H_0. Figures 5 and 6 show that the greater the absolute value of the negative skewness, the greater the default probability. Figure 5 depicts the case when the negative (left) tail freely grows to the default boundary $DL = 0$. In Figure 11, the tail of the distribution $FBP(x, t; H_0 = 1.2)$ has reached the boundary at time $t \approx 7$. The default boundary stops the tail left development and pushes the distribution to the right, changing the sign of skewness development from negative to positive. This change is accompanied by growth of the default probability (see Figures 6 and 12).

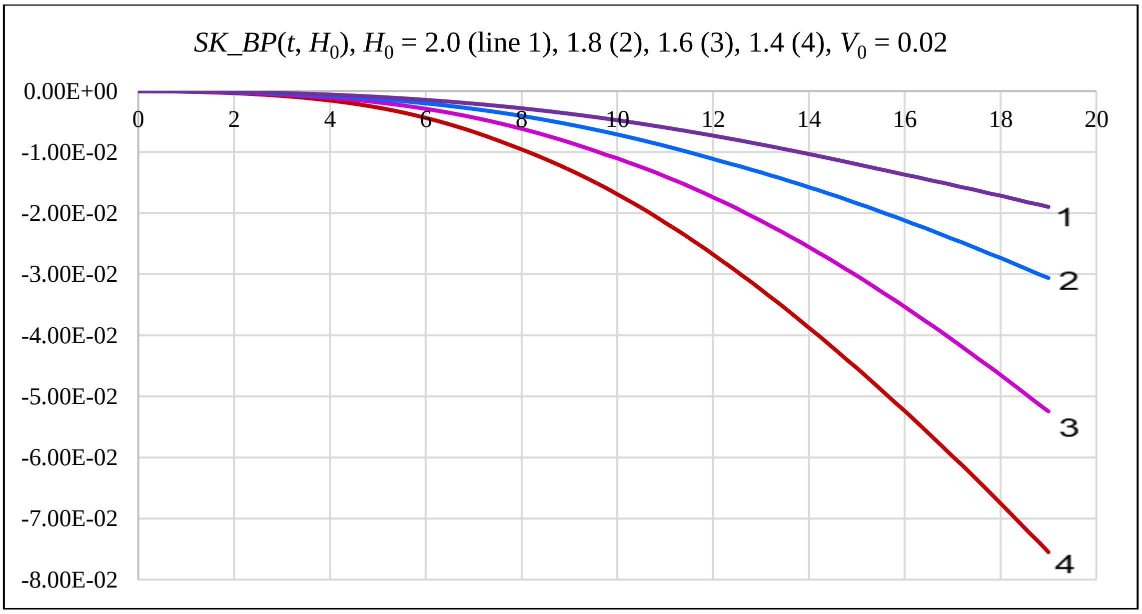

Figure 5. Skewness $SK\_BP(t)$ as a function of time (years) for various initial conditions $H_0 = 2.0$ (line 1), 1.8 (line 2), 1.6 (line 3), and 1.4 (line 4); $VAR_0 = 0.02$.

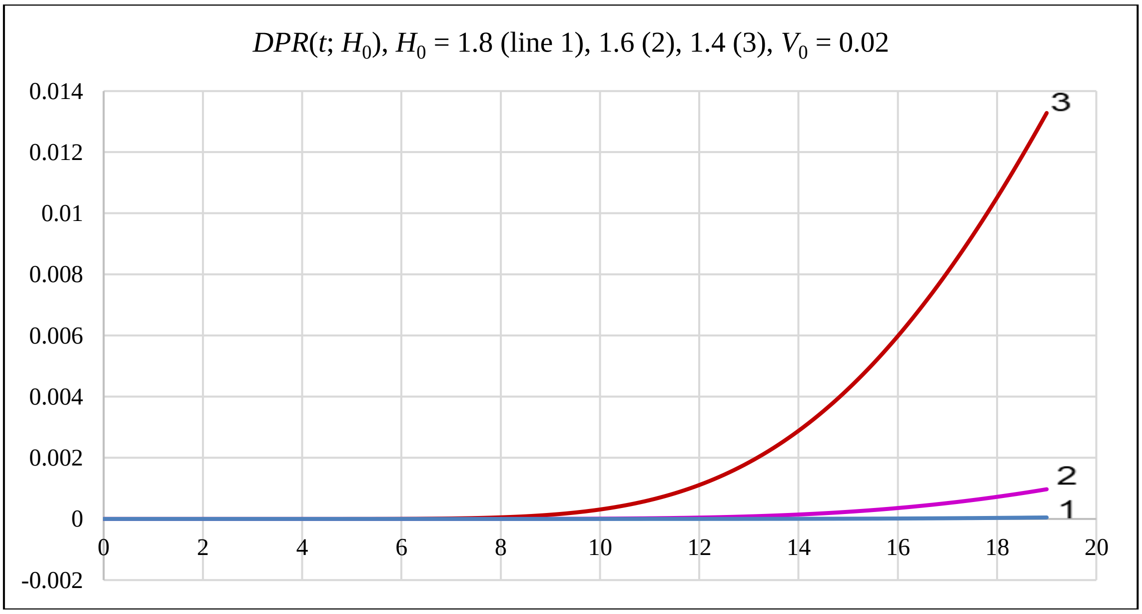

Figure 6. Default probability $DPR(t; H_0)$ as a function of time (years) for various initial conditions $H_0 = 1.8$ (line 1), 1.6 (line 2), and 1.4 (line 3); $VAR_0 = 0.02$.

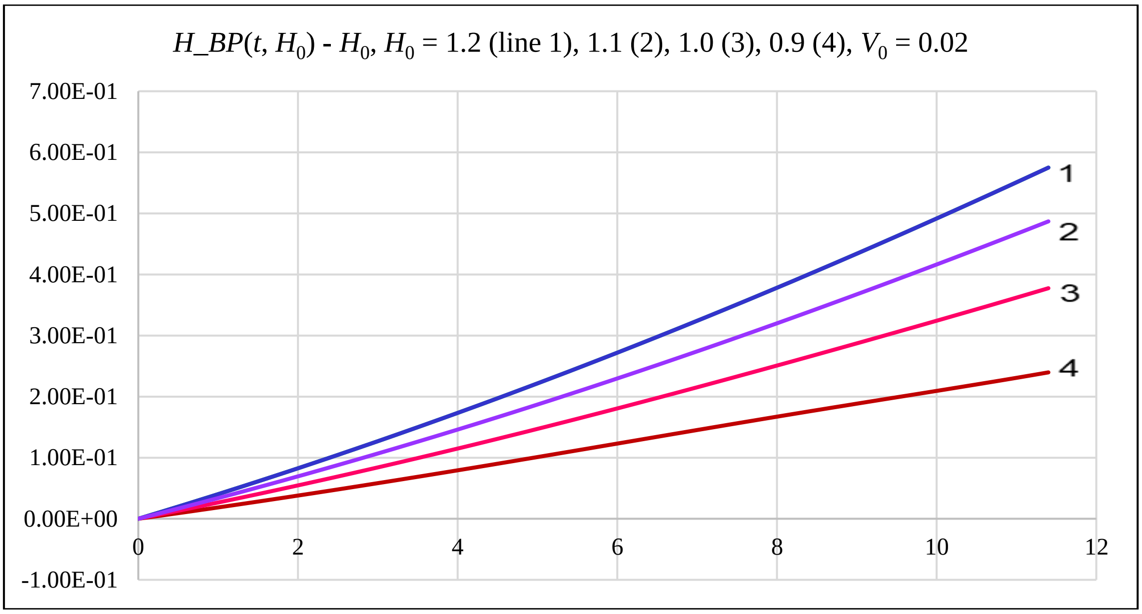

Figure 7. Mean log returns $H\_BP(t)$ - H_0 as a function of time (years) for various initial conditions H_0 = 1.2 (line 1), 1.1 (line 2), 1.0 (line 3), and 0.9 (line 4); VAR_0 = 0.02.

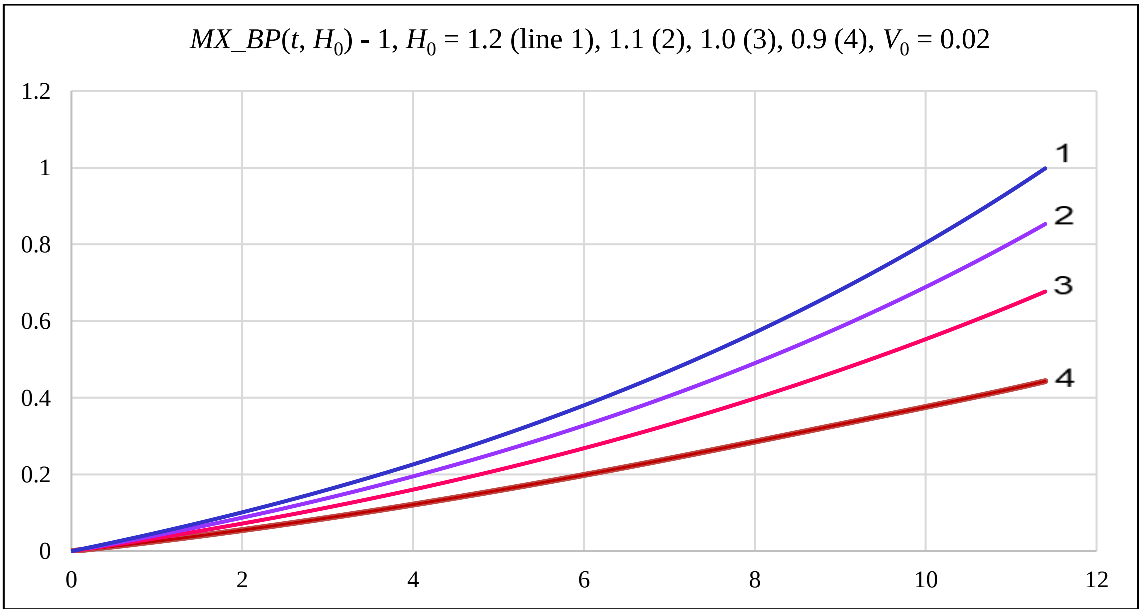

Figure. 8 Mean non-dimensional returns $MX\_BP(t)$ - 1 as a function of time (years) for various initial conditions H_0 = 1.2 (line 1), 1.1 (line 2), 1.0 (line 3), and 0.9 (line 4); VAR_0 = 0.02.

Figures 9 and 12 support the statement that for skewed return distributions of the firms paying their BSEs, the variance is not an adequate measure of risk. To optimize portfolio allocation, one must know the default probability (the intensity of default probability) as a function of time and parameters of the firm and its business environment and keep it at or below the threshold value while maximizing the mean asset returns and skewness.

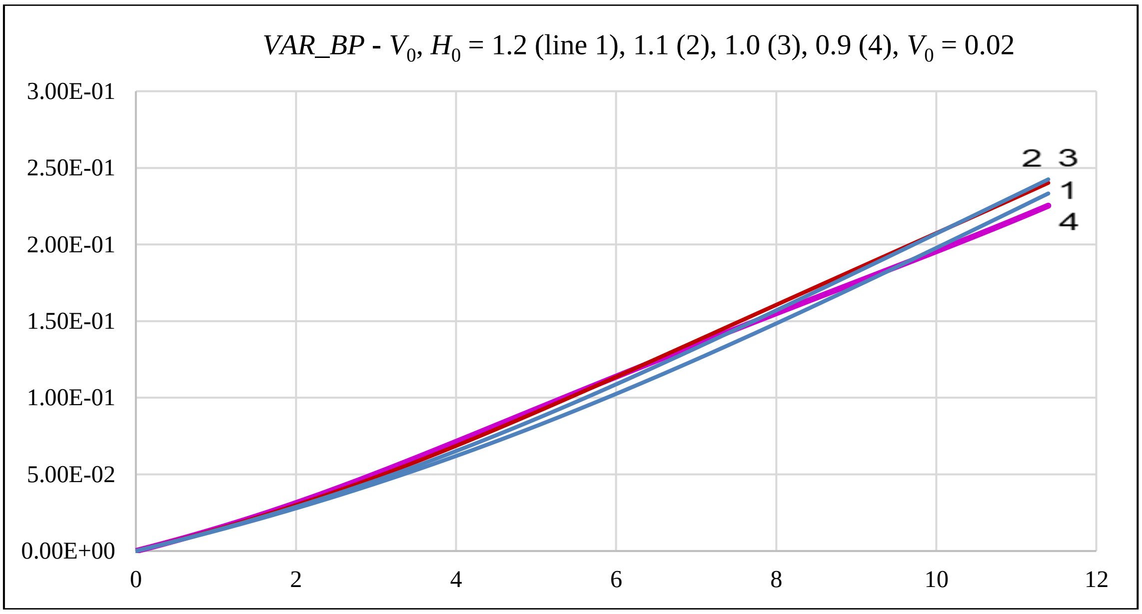

Figure 9. Variance *VAR_BP*(*t*) - VAR_0 as a function of time (years) for various initial conditions H_0 = 2.0 (line 1), 1.8 (line 2), 1.6 (line 3), and 1.4 (line 4); VAR_0 = 0.02.

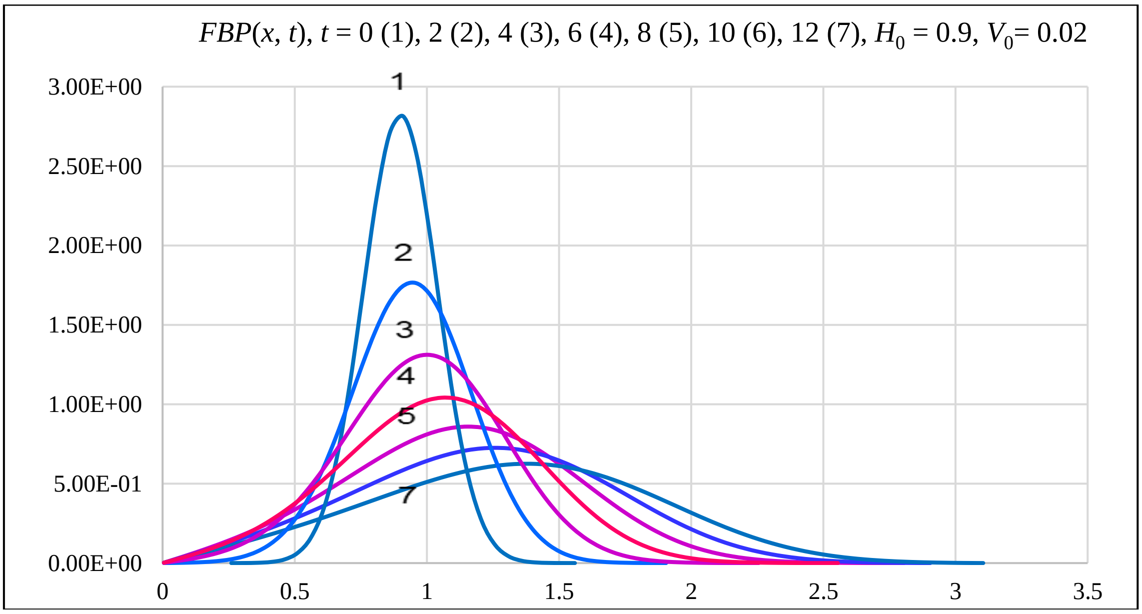

Figure 10. Evolution of the boundary problem distribution *FBP*(*x*, *t*) for H_0 = 0.9 and *t* = 0 (line 1), 2 (line 2), 4 (line 3), 6 (line 4), 8 (line 5), 10 (line 6), 12 (line 7); VAR_0 = 0.02.

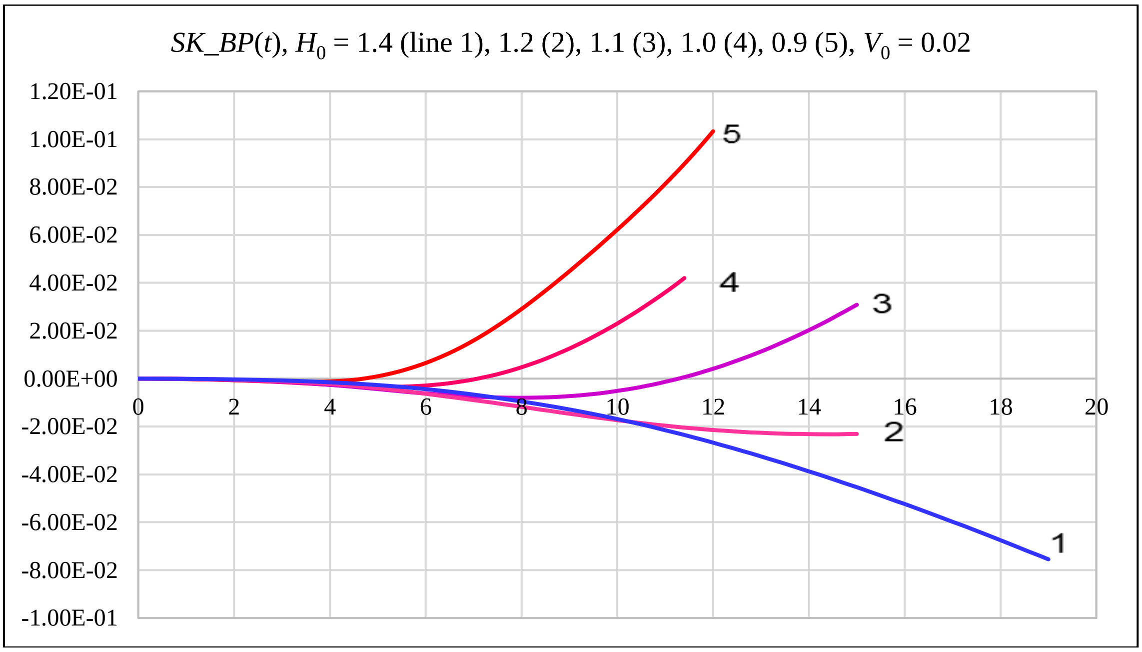

Figure 11. Skewness *SK_BP*(t) as a function of time (years) for various initial conditions H_0 = 1.4 (line 1), 1.2 (line 2), 1.1 (line 3), 1.0 (line 4), and 0.9 (line 5); VAR_0 = 0.02.

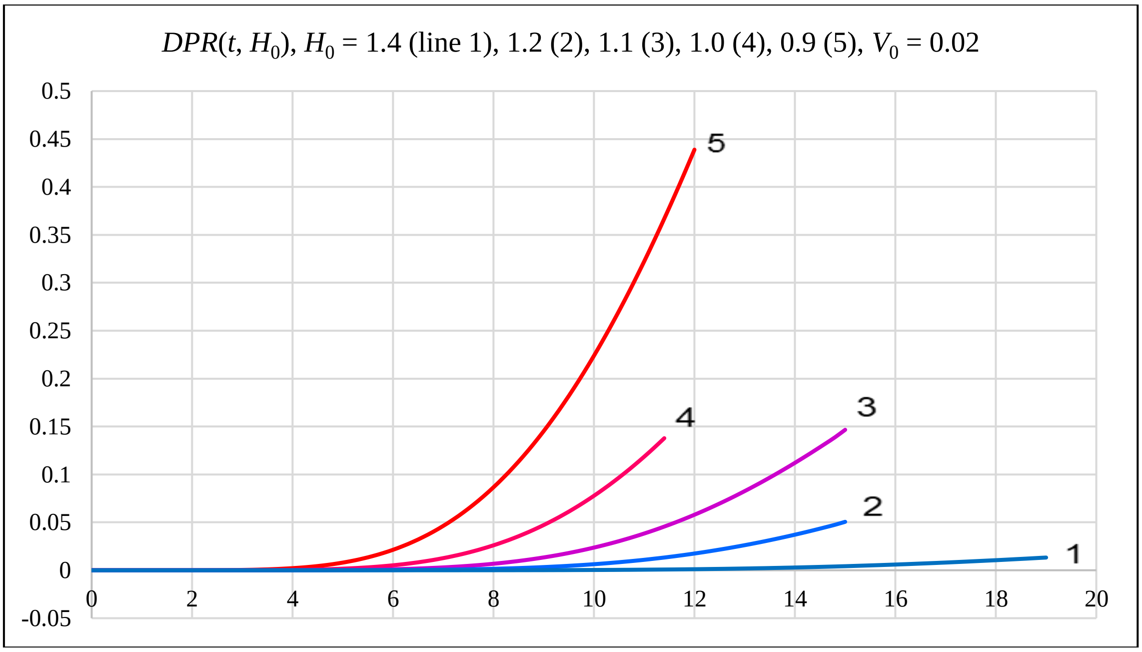

Figure 12. Default probability *DPR*(t; H_0) as a function of time (years) for various initial conditions H_0 = 1.4 (line 1), 1.2 (line 2), 1.1 (line 3), 1.0 (line 4), and 0.9 (line 5); VAR_0 = 0.02.

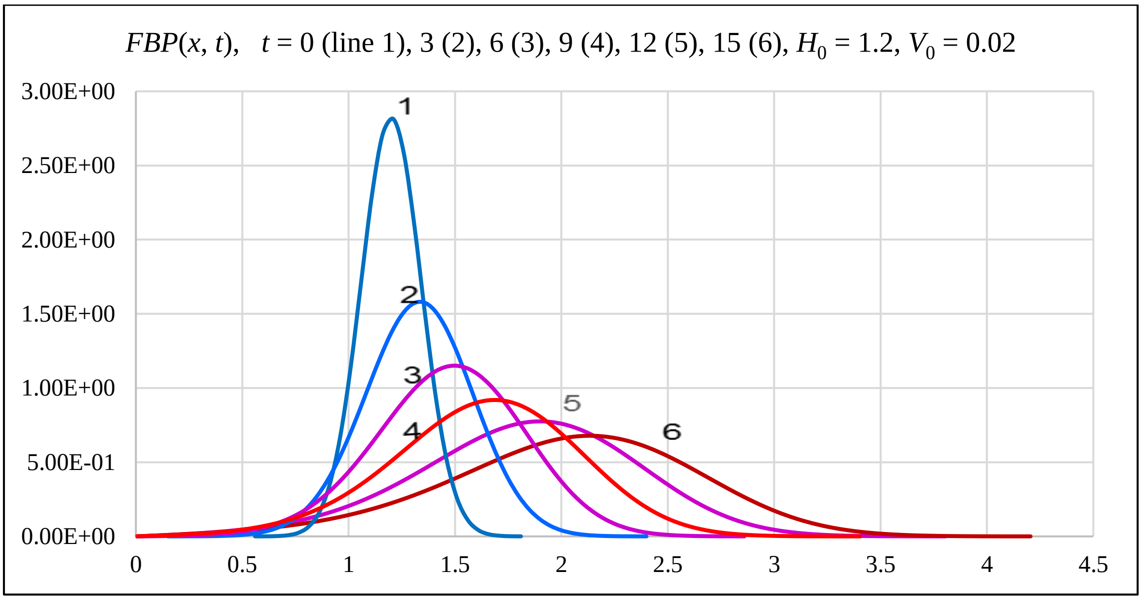

Figure 13. Evolution of the boundary problem distribution *FBP*(x, t) for H_0 = 1.2 and t = 0 (line 1), 3 (line 2), 6 (line 3), 9 (line 4), 12 (line 5), 15 (line 6); VAR_0 = 0.02.

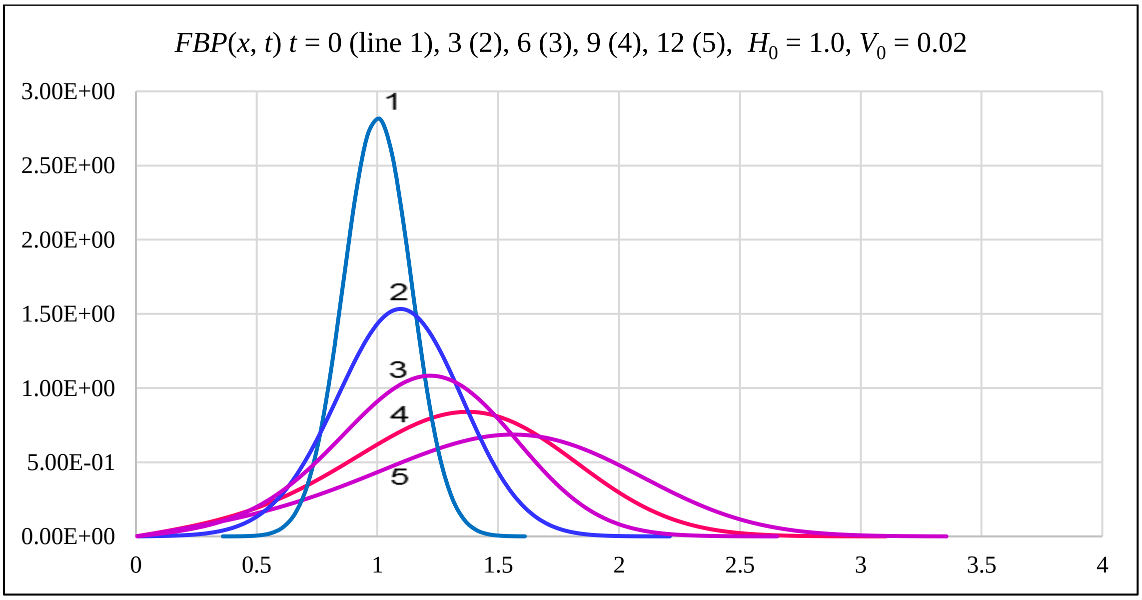

Figure 14. Evolution of the boundary problem distribution *FBP*(x, t) for H_0 = 1.0 and t = 0 (line 1), 3 (line 2), 6 (line 3), 9 (line 4), 12 (line 5); VAR_0 = 0.02.

Figures 5, 10, 13, and 14 demonstrate the development of return distributions *FBP*(x, H_0) for different initial distances to the default line H_0. Here one can see the evolution of distribution skewness from negative to positive values.

Figures 7, 8, and 12 show that the low asset growth is very deceptive: the firm assets still grow although slowly, but the default probability has achieved high values inflicting the firm high losses and threatening it with soon default.

Conclusion

The problems of asset pricing and efficient portfolio allocation have captured significant attention in the theoretical and empirical studies of investments since the middle of the 1960th. The first fundamental results in that direction are Markowitz's mean-variance optimal portfolio theory (MV optimization) and Sharp-Lintner's capital asset pricing model (CAPM). Both theories assume the normal distribution of asset returns and the lognormal distribution of assets. In the 1970th, these theories have got a support from Merton's "option hypothesis" coined as "while options are highly specialized and relatively unimportant financial instruments [. . .] the same basic approach could be applied in developing a pricing theory for corporate liabilities in general" (Merton, 1974, p. 449) because of "isomorphic correspondence between almost any corporate liability and options". Financial economists have readily accepted Merton's hypothesis and the geometric Brownian model (GBM) following from it. GBM has become the most popular model in credit risk analysis. The return distribution satisfying GBM is a normal distribution consistent with Markowitz's and Sharp-Lintner's theories, but contradicting the observed data. To analyze the results presented in literature, we use EMM taking account of the firm's business securing expenses (BSEs) and generating the return distributions whose skewness grows on time.

The empirical data accumulated in the 1980th show that the observed default frequencies are much higher than the default probabilities predicted by GBM. To meet the facts, some theorists begin to use the class of elliptic functions to describe asset returns. This class of symmetric functions includes the functions whose tails are heavier than the tails of the normal distribution (e.g., Student's *t* distribution). The researchers use elliptic functions in estimating the misspecification of asset pricing models (the Hansen-Jagannathan distance, 1997), though with little success. We explain this little success by the different shapes of efficient frontiers; a concave up surface for symmetric distributions of asset returns assumed by Hansen and Jagannathan and a surface with edges for skewed distributions (Athayde & Flores, 2004).

Meanwhile, empirical evidence of skewness of asset returns and portfolios appears, starting the period of asset pricing models and portfolio allocation methods taking account of higher moments (skewness and kurtosis) in asset returns. Some researchers studying portfolio allocation problems assume that the asset return skewness is generated by jump-diffusion processes exogenous to the firm dynamics. Because the asset return skewness is a natural consequence of the inner dynamics of the firm paying its BSEs, the results of these jump-diffusion studies are always misleading.

Another group of researchers applies skew-elliptic functions to approximate the asset return distribution. These skewed functions have a set of parameters for flexible tuning their skewness and tails. In addition to flexibility, the skew-elliptic functions allow for achieving closed-form solutions which is very helpful for solving the efficient portfolio allocation problem. To perform efficient allocation lacking information about the return distributions, the researchers must make reasonable assumptions about these distributions. Of course, optimization outcomes depend on how good those assumptions are.

A much heavier problem represents the traditional approach to portfolio optimization ascending to Markowitz's MV optimization, which, to the best of our knowledge, is used by all researchers of portfolio optimization with higher moments before the summer of 2022. When performing MV optimization, one maximizes the mean portfolio return, keeping the risks fixed. At that, the variance is considered a good measure of risk, which is absolutely true for the diffusion spread of returns. To find the efficient portfolio of skewed asset distributions in the frameworks of this model, the researchers maximize the portfolio mean return and skewness while keeping the variance fixed. However, we have shown that high positive skewness comes with a high default probability, and the unconditional maximization of portfolio skewness dangerously increases portfolio risks. Second, for skewed asset returns, the variance is no adequate measure of asset risks because the asset spread depends now on both the diffusion and growing distribution tail. In these conditions, the intensity of default probability, or the default probability itself, is a better measure of asset risks. Therefore, for higher moment portfolio optimization, one must solve the following problem: maximize the mean return and skewness while keeping the default probability under control. The EMM results presented in this paper support the thesis that the variance is no good measure of risk for the skewed distributions.

Unfortunately, the settings of the optimal portfolio allocation problem, as they are seen by all researchers studying it, do not allow estimating the default probability. We suggest a possible resolution of this conflict by joining the efforts of two approaches: EMM and the skew-elliptic function method. EMM generates the asset return distribution providing the information for tuning the skew-elliptic functions and computes the default probabilities for performing the correct portfolio optimization; the skew-elliptic function method solves the portfolio optimization problem using the data provided by EMM.

Asset pricing models using heuristic/fundamental variables meet with mixed success because of lacking knowledge about the asset return distribution and the parameters and mechanisms controlling its development. We have shown that even size SMB, the most popular variable in asset pricing models, does not uniquely determine the firm state and dynamics. The fact that the variance is no measure of risks for shaky firms means that the variance (beta) cannot be an explaining variable in the perspective asset-pricing model taking account of higher moments. A more systematic approach to the asset-pricing problem using EMM information about the asset return distributions could be helpful.

References

Aase, K. (1984). Optimal portfolio diversification in a general continuous-time model. *Stochastic Processes and Their Applications, 18*, 81-98.

Aase, K., & Oksendal B. (1988). Admissible investment strategies in continuous trading. *Stochastic Processes and Their Applications, 30*, 213-227.

Adcock, C., & Shutes, K. (2001). Portfolio selection based on the multivariate skew normal distribution. In *Financial Modelling*, Skulimowski, A. (ed.), Krakow, Progress & Business Publishers.

Adcock, C. (2002). Asset pricing and portfolio selection based on the multivariate skew-student distribution. *Annals of Operations Research, 176* (1), 221-234.

Adcock, C., Eling, M., & Loperfido, N. (2012). Skewed distributions in finance and actuarial science. *European Journal of Finance, 21*(13), 1-29.

Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic production function model. *Econometrics, 12*(1), 21-37.

Allen, D., & McAleer, M. (2018). “Choosing factors” by Fama and French (2018): A comment. Retrieved from https://ssrn.com/abstract=3272608

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Financial Economics, 63*, 443-494.

Arnold, B., & Beaver, R. (2000). Hidden truncation models. *Sankhya, Series A, 62*, 22-35.

Athayde, G., & Flores, R. (2004). Finding a maximum skewness portfolio—A general solution of three moments portfolio choice. *Economic Dynamics and Control, 28*(7), 1335-1352.

Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika, 83*, 715-726.

Azzalini, A., & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B, 65*, 367-389.

Azzalini, A. (2005). The skew-normal distribution and related multivariate families (with discussion by Marc G. Genton and a rejoinder by the author). *Scandinavian Journal of Statistics, 32*, 159-200.

Badrinath, S., & Chatterjee, S. (1988). On measuring skewness and elongation in common stock return distribution: The case of market index. *Journal of Business, 61*, 451-472.

Bae, K.-H., Karolyi, A., & Stulz, R. (2003). A new approach to measuring financial contagion. *Review of Financial Studies, 16*(3), 717-763.

Banz, R. (1981). The relationship between return and market value of common stocks. *Financial Economics, 9*(1), 3-18.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society, 353*, 401-419.

Barone-Adesi, G., Gagliardini, P., & Urga, G. (2004). Testing asset pricing models with co-skewness. *Business & Economic Statistics, 22*(4), 474-485.

Basu S. (1983). The relationship between earning yield, market value, and return for NYSE common stocks: Further evidence. *Financial Economics, 12*, 129-156.

Bhandari, L. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *Journal of Finance, 43*(2), 507-528.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Political Economy, 81*, 637-654.

Bohn, J. (2006). Structural modelling in practice. *Moody's.* Retrieved from http://web-docs.stern.nyu.edu/salomon/docs/Credit2006/Bon.pdf

Boudt, K., Cornilly D., & Verdonck, T. (2020). A co-skewness shrinkage approach for estimating the skewness of linear combinations of random variables. *Financial Econometrics, 18*(1), 1-23.

Branco, M., & Dey, D. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis, 79*, 99-113.

Briec, W., Kerstens, K., & Lesourd, J. (2004). Single-period markowitz portfolio selection, performance gauging, and duality: A variation on the Luenberger shortage function. *Optimization Theory and Applications, 120*(1), 1-27.

Briec, W., Kerstens, K., & Jokung, O. (2007). Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management Science, 53*(1), 135-149.

Briec, W., & Kerstens, K. (2010). Portfolio selection in multidimensional general and partial moment space. *Economic Dynamics and Control, 34*, 636-656.

Brown, G., & Matysiak, G. (2000). Real estate investment: A capital market approach. *Financial Times*, Prentice Hall, London.

Chen, N., & Kou, S. (2009). Credit spreads, optimal capital structure, and implied volatility with endogenous default and jump risks. *Mathematical Finance, 19*(3), 343-378.

Cox, J., Ross, S., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Financial Economics, 7*(3), 229-263.

Cvitanic, J., Polimenis, V., & Zapatero, F. (2008). Optimal portfolio allocation with higher moments. *Annals of Finance, 4*, 1-28.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross-sectional variation in stock returns. *Journal of Finance, 52*(1), 1-33.

Das, S., & Uppal, R. (2004). Systemic risk and international portfolio choice. *Journal of Finance, 59*, 2809-2834.

Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference and evidence from the cross-section of equity returns. *Journal of Finance, 57*, 369-403.

Drew, M., & Veeraraghavan, M. (2002). A closer look at the size and value premium in emerging markets: Evidence from the Kuala Lumpur stock exchange. *The Asia Pacific Economy, 8*(3), 354-379.

Durga, S. (2020). Examining the validity of three-factor Fama and French model in Indian stock markets: An empirical analysis. *International Journal of Management, 11*(8), 89-94.

Edirisinghe, N., & Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Banking and Finance, 31*, 3311-3335.

Engle, R., Ito, T., & Lin, W-L. (1990). Meteor showers or heat waves? Heteroskedastic intra-day volatility in the foreign exchange market. *Econometrica, 58*(3), 525-542.

Eraslan, V. (2013). Fama and French three-factor model: Evidence from Istanbul stock exchange. *Business and Economic Research, 4*(2), 11-22.

Erb, C., Harvey, C., & Viscanta, T. (1994). Forecasting international equity correlations. *Financial Analyst Journal, 50*(1), 32-45.

Fama, E., & French, K. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427-465.

Fama, E., & French, K. (1993). Common risk factors in the returns of stocks and bonds. *Financial Economics, 33*(1), 3-56.

Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55-84.

Fama, E., & French, K. (2004). The capital asset pricing model: Theory and evidence. *Economic Perspectives, 18*(3), 25-46.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Financial Economics, 116*(1), 1-25.

Fama, E., & French, K. (2018). Choosing factors. *Financial Economics, 128*(2), 234-252.

Fang, H., & Lai, T. (1997). Co-kurtosis and capital asset pricing. *Financial Review, 32*(2), 293-307.

Friend, I., & Westerfield, R. (1980). Co-skewness and capital asset pricing. *Journal of Finance, 35*, 897-913.

Gospodinov, N., Robotti, C., & Kan, R. (2012). On the Hansen-Jagannathan distance with no-arbitrage constrain. Retrieved from SSRN: https://ssrn.com/abstract=1534727

Gospodinov, N., Robotti, C., & Kan, R. (2013). Chi-squared tests for evaluation and comparison of asset pricing models. *Econometrics, 173*(1), 108-125.

Gospodinov, N., Robotti, C., & Kan, R. (2014). Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *Review of Financial Studies, 27*(7), 2139-2170.

Hansen, L., & Jagannathan R. (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance, 52*(2), 557-590.

Harrison, J., & Kreps, D. (1979). Martingale and arbitrage in multiperiod securities markets. *Economic Theory, 20*, 381-408.

Harrison, J., & Pliska, R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications, 11*, 215-260.

Harvey, C., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance, 55*(3), 1263-1295.

Harvey, C., Liechty, J., Liechty, M., & Mueller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance, 10*(5), 469-485.

Harvey, C., Liu, Y., & Zhu, H. (2016). …and the cross-section of expected returns. *Review of Financial Studies, 29*(1), 5-68.

Hilberink, B., & Rogers, L. (2002). Optimal capital structure and endogenous default. *Finance and Stochastics, 6*(2), 237-263.

Huang, J., & Huang, M. (2002, revised May 2003). How much of the corporate-treasury yield spread is due to credit risk? Results from a new calibration approach. Working Paper, GSB, Stanford University.

Hugonnier, J., Malamud, S., & Morellec, E. (2015). Credit market frictions and capital structure dynamics. *Economic Theory, 157*(C), 1130-1158.

Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance, 51*(1), 3-53.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*(1), 65-91.

Jondeau, E., & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management, 12*(1), 29-55.

Kou, S. (2002). A jump-diffusion model for option pricing. *Management Science, 48*, 1086-1101.

Kraus, A., & Litzenberger, R. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance, 31*(4), 1085-1100.

Krueger, J. (2021). Nonparametric portfolio efficiency management with higher moments. *Empirical Economics, 61*, 1435-1469.

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Empirical Finance, 10*, 603-621.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Multivariate Analysis, 88*(2), 365-411.

Leland, H. (1994). Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance, 49*, 1213-1252.

Leland, H., & Toft, K. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance, 51*, 987-1019.

Lettau, M., & Ludvigson, S. (2001). Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Political Economy*, *109*(6), 1238-1287.

Lewellen, J., & Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Financial Economics, 83*, 289-314.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13-37.

Liu, J., Longstaff, F., & Pan, J. (2003). Dynamic asset allocation with event risk. *Journal of Finance, 58*, 231-259.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance, 56*(2), 649-676.

Madan, D., & Milne, F. (1991). Option pricing with VG martingale components. *Mathematical Finance*, *1*, 39-55.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77-91.

Martinelli L., & Ziemann, V. (2010). Improved estimates of higher order components and implications for portfolio selection. *Review of Financial Studies, 23*, 1467-1502.

Mencia, J., & Sentana, E. (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *Econometrics, 153*(2), 105-121.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Economic Theory, 3*(4), 373-413.

Merton, R. (1974). On the pricing of corporate debt: Risk structure of interest rates. *Journal of Finance, 29*, 449-470.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Financial Economics, 3*(1-2), 125-144.

Morey, M., & Morey, R. (1999). Mutual fund performance appraisals: A multi-horizon perspective with endogenous benchmarking. *Omega, 27*(2), 241-258.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768-783.

Patton, A. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Financial Econometrics, 2*(1), 130-168.

Richardson, M., & Smith, T. (1993). A test for multivariate normality in stock returns. *Journal of Business, 66*, 295-321.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Portfolio Management, 11*(3), 9-16.

Sahu, S., Branco, M., & Dey, D. (2003). A new class of multivariate skew distributions with application to Bayesian regression models. *Canadian Journal of Statistics, 12*(1), 129-150.

Sentana, E. (2009). The econometrics of mean-variance efficiency tests: A survey. *Econometrics, 12*, C65-C101.

Shanken, J. (1987). Multivariate proxies and asset prices relations. Living with the roll critique. *Financial Economics, 18*, 91-110.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425-442.

Shapovalova, K., Subbotin, A., & Chauveau, T. (2011). Returns premia on company fundamentals. Retrieved from https://ssrn.com/abstract=1770744

Shemetov, V. (2020a). Continuous-time models for firm valuation and their collateral effect on risk-neutral probabilities and no-arbitraging principle. *Management Studies, 8*(3), 191-214.

Shemetov, V. (2020b). Credit risk model taking account of inflation and its contribution to macroeconomic discussion on effect of inflation on output growth. *Management Studies, 8*(6), 430-452.

Shemetov, V. (2021). Effect of corporate debt on firm value. *Management Studies, 9*(5), 368-411.

Simkowitz, M., & Beedies, W. (1980). Asymmetric stable distributed security returns. *Journal of American Statistical Association, 75*, 306-312.

Singleton, J., & Wingender, J. (1986). Skewness persistence in common stock returns. *Financial and Quantitative Analysis, 21*, 335-341.

Stattman, D. (1980). Book values and stock returns. *Journal of Selected Papers, 4*, 25-45.

Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *Journal of Finance, 59*, 831-868.

Zhou, C. (2001). The term structure of credit spreads with jump risk. *Banking and Finance, 25*(11), 2015-2040.