RBGN REVISTA BRASILEIRA DE GESTÃO DE NEGÓCIOS
© FECAP

ISSN 1806-4892
e-ISSN 1983-0807

OPEN ACCESS

ARTICLE

Asset pricing: an alternative estimation for the five-factor model

Renan Oliveira Regis[1]
Raydonal Ospina[2,3]
Wilton Bernardino[4]

Received on:
Feb/02/2023
Approved on:
Sept/10/2024

Responsible editor:
Prof. Dr. Joelson Sampaio

Reviewers:
Carlos Eduardo Lavarda;
Gabriel Carvalho

Evaluation process:
Double Blind Review

This article is open data

Abstract

Purpose – This paper introduces a comprehensive approach to estimating the five-factor model in financial markets, emphasizing flexibility and predictive improvement via GAMLSS models. We highlight the innovative potential of this methodology in asset pricing theory.

Theoretical framework – This paper seeks to evaluate the behavior of asset prices under conditions of uncertainty. Fama and French (2015) inspired us to present an extension via structured additive distributional regression using GAMLSS for the five-factor model.

Design/methodology/approach – The sample contains information from the Brazilian financial market from 1994 to 2018. Given the violation of the conditional normal distribution commonly observed in these data, we propose adopting GAMLSS modeling. This approach allows for the flexibility of probability distributions associated with stock portfolio returns, more accurately accommodating location and scale.

Findings – GAMLSS modeling significantly enhances predictive performance, providing a robust alternative to traditional models that use the normal distribution. Furthermore, no evidence of specification error was observed using GAMLSS models, reinforcing their reliability.

Practical & social implications of research – The use of flexible GAMLSS modeling for asset pricing is proposed In the Brazilian financial market. This would improve decision-making capacity related to financial markets and asset pricing.

Originality/value – In terms of contribution, the article proposes a new estimation approach for the five-factor model using GAMLSS models.

Keywords: Asset pricing, five factors, GAMLSS, Fama and French.

1. Universidade Federal de Pernambuco, Departamento de Economia, Recife, PE, Brasil
2. Universidade Federal da Bahia, Departamento de Estatística, Salvador, BA, Brasil
3. Universidade Federal de Pernambuco, Departamento de Estatística, Recife, PE, Brasil
4. Universidade Federal de Pernambuco, Departamento de Ciências Contábeis e Atuariais, Recife, PE, Brasil

How to cite:
Regis, R. O., Ospina, R., & Bernardino, W. (2024). Asset pricing: an alternative estimation for the five-factor model. *Revista Brasileira de Gestão de Negócios*, *26*(3), e20230205. https://doi.org/10.7819/rbgn.v26i03.4272

FUNDAÇÃO ÁLVARES PENTEADO
FECAP

Revista Brasileira de Gestão de Negócios

https://doi.org/10.7819/rbgn.v26i03.4272

1 Introduction

Asset pricing theory explores how risk factors influence financial market prices under uncertainty. In this sense, returns are derived from risk premiums that investors demand to compensate for the risk of their investments (Regis et al., 2023). The main objective of the theory is therefore to explain the determination of asset prices.

The Capital Asset Pricing Model (CAPM) pioneered a model based on a linear relationship between the expected return on an asset and the market risk premium. The CAPM makes it possible to estimate the rate of return (β) taking into account the systematic risk of the market (Sharpe, 1964). However, Miller and Scholes (1972) strongly criticized it, pointing out that the estimator of the β parameter is often biased due to the difficulty of finding a single response variable that is representative of the financial market.

The Price Arbitrage Theory (APT), proposed by Ross (1976), allows the analysis of risk factors that influence the return on assets, which can be macroeconomic or market-specific. This theory underpins the transition to multifactor models, such as those of Fama and French (1993, 2015), which add specific characteristics of companies to the market risk predicted by the CAPM.

According to Fama and French (1993), stock returns are affected by three risk factors. The first is general market risk, represented by the market premium in the CAPM model. The other two factors, called size (market value) and book-to-market (ratio of equity to market value), are intrinsic characteristics of companies. Fama and French (1993) conclude that the market premium alone contains little information about asset returns, while size and book-to-market capture significant variations in returns.

Fama and French (2015) updated the model proposed in 1993 to include the risk premiums associated with profitability and investment, creating the five-factor model. The authors analyzed stock data from the NYSE, NASDAQ and Amex between 1963 and 2013 and concluded that the inclusion of these premiums makes the book-to-market premium redundant in explaining returns. By introducing a transformation in the book-to-market factor, renaming it the orthogonalized factor, they showed that the five-factor model provides a more accurate analysis of returns than the three-factor model.

Leite et al. (2020) find that variables associated with future investment opportunities impact the excess returns of assets, but are not correlated with Fama and French's (2015) five factors. By incorporating exogenous shocks to macroeconomic factors, the authors conclude that shocks to inflation and the term structure of interest rates, for example, explain average returns better than in the usual five-factor model. On the other hand, Carvalho et al. (2021) analyze data from emerging markets in Latin America and find that Fama and French's (2015) five-factor model has excellent predictive performance. This study provides evidence that the five-factor model can also be applied to growing emerging economies.

Asset pricing models typically employ an estimation procedure such as ordinary least squares (OLS), maximum likelihood and/or robust techniques with heteroskedasticity correction, taking into account distributional conditions associated with the normality of errors. However, several studies, including Carvalhal & Mendes (2003), Rocco (2014) and Regis et al. (2023), show that extreme values in stock returns occur more frequently than expected under the assumption of normality.

Regis et al. (2023) found evidence of non-normality in data from the Brazilian financial market and used the Generalized Additive Models for Location, Scale and Shape (GAMLSS) in asset pricing (Rigby & Stasinopoulos, 2005). In this sense, the authors were motivated by the flexibility of this statistical model in dealing with different forms of distributions associated with the data under analysis. However, the study highlights that the intention to create 75 portfolios led to situations in which some of them were not formed, while others were composed with a very small number of assets.

This article adopts a different approach to the study conducted by Regis et al. (2023) by adjusting the portfolio formation strategy to include 18 of them in order to ensure more comprehensive diversification. This adjustment implies a significant increase in the number of assets included in each portfolio. The results obtained revealed patterns that indicate greater similarity between the Brazilian financial market and the American market, including the redundancy of the book-to-market factor.

The adoption of GAMLSS modeling takes into account the empirical distribution of financial market data, seeking to correct the specification error identified in the response variable (excess return) when using models that assume distributional normality. The emphasis on considering the empirical distribution not only provides a more accurate representation of the reality of financial data, but also contributes to the robustness and validity of the results obtained. Therefore, the research offers a contribution to the advancement of the literature in terms of the modeling technique discussed in Regis et al. (2023), providing an analysis that considers a greater degree of diversification between portfolios, allowing us to find relationship patterns between risk premiums and excess returns not identified in the aforementioned research.

Defining excess return as the difference between the return on the asset with risk of loss and the return on the risk-free asset, the main objective is to present a new estimation approach for the five-factor model using GAMLSS models and to show that this approach improves the analysis and explanatory power of the model. This corrects the specification error present in the methods traditionally used. The study compares the results of the estimates obtained when applying the model in Brazil with the results found by Fama and French (2015), using American data.

To make the results comparable with the existing literature, we adopted the GAMLSS estimates for the mean with a linear predictor, as in the case of the linear model, but with distributional flexibility and keeping the scale and shape parameters constant. Subsequently, we compared the results of the GAMLSS modeling and the standard linear modeling through performance evaluations using information criteria such as the global deviation (GD), the AIC (Akaike information criterion), and the SBC (Bayesian information criterion). The choice of GAMLSS aims to improve the fit of the model and overcome the limitations of methods that assume a Gaussian distribution. This alternative approach provides flexibility in choosing the best distributions to fit the Brazilian financial market data.

In addition to this introduction, this article is structured as follows. Section 2 provides information on the five-factor model. Section 3 describes the methodology, with an emphasis on GAMLSS modeling. Section 4 presents the results of the applications to the Brazilian financial market. Finally, Section 5 provides concluding remarks.

2 Fama and French's five-factor models

Fama and French (1993) introduced a three-factor model that added two new factors to the model with only the market premium: size (MV), measured by market value (share price times total shares outstanding), and the book-to-market ratio (ratio of book value to market value). Fama and French (2015) proposed an extension of the three-factor model to include two additional regressors (the profitability (OP) and investment (INV) risk premiums). Further studies include Fama and French (2016), Carvalho et al. (2021) and Regis et al. (2023).

Understanding the five-factor model is based on defining the regressor variables associated with company financial/accounting data. This defines small caps and large caps (companies with a lower market value versus those with a higher value). A positive risk premium is expected for stocks of small-cap companies, as they are largely start-ups with potential for future growth. The size risk premium is calculated on the basis of the market value data at the close of the balance sheet in December of each previous year. The book-to-market ratio is calculated by dividing equity by market value (NE/MV). This indicator reveals whether the shares are overvalued (MV>NE) or undervalued (NE>MV) and reflects the comparison between the real value of the company and its speculation on the stock market. A positive premium is expected for undervalued companies. The closing balance sheet data for December of the previous year is used to calculate equity and market value, thus forming the book-to-market risk premium. Profitability is determined by dividing earnings before interest and taxes (EBIT) at the end of the previous year by net equity in the same period (EBIT/NE), thus forming the profitability premium. It is expected that higher risk premiums are associated with more profitable companies, especially since part of the profits are distributed in the form of dividends.

The investment factor signals the growth of assets on the balance sheet,

$$\frac{TotalAssets_{t-1} - TotalAssets_{t-2}}{TotalAssets_{t-2}}. \quad (1)$$

In Equation (1) $TotalAssets_{t-1}$ refers to the closing of the balance sheet of the previous year and $TotalAssets_{t-2}$ refers to the second previous year. Companies with high levels of investment are expected to compromise their ability to generate profits in the short term, resulting in lower premiums for companies with high levels of investment compared to those with low levels.

The five-factor risk premiums are SMB (small minus big), HML (high minus low NE/MV), RMW (robust minus weak OP) and CMA (conservative minus aggressive INV), respectively, as described in Fama & French (2015). The SMB premium is calculated from the difference between the returns of smaller and bigger companies. HML reflects the difference between the returns of companies with high and low book values. RMW measures the difference in returns between companies with high (robust) profits and companies with low (weak) profits. Finally, CMA measures the difference in returns between companies with conservative investments and those with aggressive investments. The criteria that determine the values of these premiums are described in Section 2.2.

The risk premiums associated with each of the five factors are calculated based on the portfolios constructed from the monthly returns. In the linear regression model, the equation is given by

$$R_{i,t} - R_{f,t} = a_i + b_i\left(R_{m,t} - R_{f,t}\right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{i,t}, \quad (2)$$

In Equation (2) $R_{i,t}$ represents the return of portfolio i in period t; $R_{f,t}$ defines the return of the risk-free asset in period t; $R_{m,t}$ denotes the return of the market portfolio in period t; SMB_t, HML_t, RMW_t and CMA_t express the risk premiums for the size, NE/MV, profitability (OP) and investment (INV) factors in period t, respectively; and $e_{i,t}$ denotes the error term of the model. According to Fama and French (2015), assuming that the parameters in (2) represent the true values, and that b_i, s_i, h_i, r_i and c_i capture all the variation in expected returns, the intercept a_i should be statistically equal to zero for all stock portfolios.

2.1 Portfolios under analysis for the Brazilian case

Due to the particularities of the Brazilian financial market, it is necessary to adapt the portfolios proposed by Fama and French (2015) for application in Brazil. According to Vieira et al. (2017), the limited size of the Brazilian market makes it difficult to build diversified portfolios. The small number of companies listed on the B3, combined with the low trading volume, prevents the replication of the portfolios outlined by Fama and French (2015).

The factor variables are organized in ascending order and segmented into specific groups. The size risk factor, represented by the market value (MV) of companies, is divided into two groups using the MV median: companies below the median are classified as "small" (S) and those above as "big" (B). Similarly, for the book-to-market factor, calculated as equity divided by market value (NE/MV), companies are grouped as "low" (L) for values below the 30% percentile, "high" (H) for values above the 70% percentile, and "neutral" (N) for intermediate values. Similarly, for profitability (OP) and investment (INV), companies are categorized as "weak" (W) or "conservative" (C) if they are below the 30% percentile, "robust" (R) or "aggressive" (A) if they are above the 70% percentile, and "neutral" (N) if they are in between.

After stratifying the variables, portfolios were formed by the intersections of the groups formed. The two MV groups were combined with the three groups of the other variables to create 2x3 portfolios, resulting in six portfolios for each combination of MV and NE/MV, MV and OP, and MV and INV, as shown in Table 1. The portfolios were made up of monthly excess returns, i.e. the average return of the portfolios (weighted by market value) minus the return of the risk-free asset, using the CDI rate in Brazil as a reference.

Table 1
2×3 portfolio formation

Portfolios $VM \times NE / MV$		NE / MV		
		Low (L)	**Neutral (N)**	**High (H)**
MV	Small (S)	SL	SN	SH
	Big (B)	BL	BN	BH
Portfolios $MV \times OP$		OP		
		Weak (W)	Neutral (N)	Robust (R)
MV	Small (S)	SW	SN	SR
	Big (B)	BW	BN	BR
Portfolios $MV \times INV$		INV		
		Conservative (C)	Neutral (N)	Aggressive (A)
MV	Small (S)	SC	SN	SA
	Big (B)	BC	BN	BA

According to Fama and French (2015), portfolios should be rebalanced annually to capture market dynamics. In the analysis, the portfolios are rebalanced in January, based on the consolidated balance sheet of the companies in December of the previous year, which requires two years to form the investment variable. Further details on the criteria used to select the companies are described in the supplementary material.

2.2 Construction of the five-factor risk premiums in the Brazilian case

The market risk premium is based on the IBOVESPA, representing the performance of the Brazilian stock market. The risk premiums for the size, book-to-market, profitability and investment factors are divided into groups, as shown in Table 2. Thus, the portfolios are formed by combining these groups, labeled by size (S or B for small or big) and by characteristics such as book-to-market (H, N, L for high, neutral, low), profitability (R, N, W for robust, neutral, weak) and investment (C, N, A for conservative, neutral, aggressive). Fama and French (2015) adopt a 2×3 portfolio structure, crossing the size factor with the other factors.

Consolidated data from the companies' annual balance sheets are used to calculate the risk premiums. The portfolios are fixed and the risk premiums are calculated monthly, repeating the process each year. The portfolios are then rebalanced annually, always in January, to produce the risk premiums. Recent work follows the same strategy when defining the calendar and rebalancing portfolios, considering time windows between January and December for variable income asset data in Brazil (see Regis et al., 2023 and Silva Jr, 2023).

3 Methodology

In many applications, it is critical to capture the uncertainties of predictions beyond mere point estimates. There are several approaches for making inferences about random aspects and estimating predictive distributions. Examples include quantile regression (Koenker, 2005) or Structured Additive Distributional Regression (SADR) (Klein et al., 2015). SADR relates several potentially different additive predictors to each parameter of an arbitrary parametric reference distribution. This allows the practitioner to explicitly model the random aspect of the data generation processes, and thus learn about the entire distribution.

3.1 Generalized additive models for location, scale and shape

Generalized Additive Models for Location, Scale and Shape (GAMLSS) are a class of SARD models that make the assumption that the response variable follows a specific parametric probability distribution function. GAMLSS models (Rigby & Stasinopoulos, 2005) extend linear models and Generalized Linear Models (GLMs) (Nelder & Wedderburn, 1972), as well as Generalized Additive Models (GAMs) (Hastie & Tibshirani, 1995), which allow for structured additive predictors.

The modeling strategy allows for the consideration of probability distributions with highly flexible shapes,

Table 2
Construction of risk premiums for factors

$SMB_{\frac{NE}{MV}}$	$(SH+SN+SL)/3-(BH+BN+BL)/3$
SMB_{OP}	$(SR+SN+SW)/3-(BR+BN+BW)/3$
SMB_{INV}	$(SC+SN+SA)/3-(BC+BN+BA)/3$
SMB	$(SMB_{NE/MV}+SMB_{OP}+SMB_{INV})/3$
HML	$(SH+BH)/2-(SL+BL)/2=\left[(SH-SL)+(BH-BL)\right]/2$
RMW	$(SR+BR)/2-(SW+BW)/2=\left[(SR-SW)+(BR-BW)\right]/2$
CMA	$(SC+BC)/2-(SA+BA)/2=\left[(SC-SA)+(BC-BA)\right]/2$

Note. Source: Retrieved from "A five-factor asset pricing model," Fama and French, 2015, p. 6.

including continuous or discrete patterns with varying degrees of skewness and kurtosis. The model's regressors can be extended to model additional parameters of the response variable distribution in addition to the mean. This approach is valuable when dealing with response variables (such as financial data) that do not follow an exponential family distribution, such as the Gaussian distribution (Rigby & Stasinopoulos, 2005).

In the GAMLSS model, the observations y_t of the response variable are assumed to be independent for $t = 1, 2, \ldots, T$, and follow the probability density function $f\left(y_t \mid \theta^t\right)$ conditioned on the information contained in the parameter vector $\theta^t = \left(\theta_{1t}, \theta_{2t}, \theta_{3t}, \theta_{4t}\right) = \left(\mu_t, \sigma_t, \nu_t, \tau_t\right)$. Each of the components of the parameter vector can be modeled as a function of some semi-parametric predictor. The parameters μ_t and σ_t generally represent location (e.g., mean or median) and scale (e.g., precision or variance), while the other parameters, such as (ν_t, τ_t), can capture characteristics such as asymmetry and kurtosis. Considering a vector of observations of the response variable of size T, represented in this article by the excess returns of a portfolio in each period considered, we can express $y^\top = (y_1, y_2, \ldots, y_T)$. For $k = 1, 2, 3, 4$, each component of the parameter vector is modeled by a predictor using a known link function $g_k(\cdot)$, i.e. the systematic component (submodel) is of the form

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{j_k} Z_{jk} \gamma_{jk}. \tag{3}$$

In Equation 3, the vector θ_k can be composed of the parameters μ, σ, ν and τ, which are vectors of dimension T, $\beta_k^\top = \left(\beta_{1k}, \beta_{2k}, \ldots, \beta_{j'_k k}\right)$ is the vector of size fixed effect parameters j'_k, X_k is the design matrix (covariate matrix) with known variables of order $T \times j'_k$ for the fixed effects (in this paper, this matrix is composed of the risk premiums of the five factors), Z_{jk} is a known design matrix for the random effects of order $T \times q_{jk}$, and γ_{jk} is a q-dimensional random effects variable q_{jk}, which is distributed as $\gamma_{jk} \sim N_{q_{jk}}\left(0, G_{jk}^{-1}\right)$, where G_{jk}^{-1} is the inverse of the symmetric matrix $G_{jk} = G_{jk}\left(\lambda_{jk}\right)$ of size $q_{jk} \times q_{jk}$, which depends on a vector of hyperparameters λ_{jk}. If G_{jk} is singular, then γ_{jk} has a density function proportional to $exp\left(\frac{-1}{2} \gamma_{jk}^t G_{jk} \gamma_{jk}\right)$.

The flexibility of Equation 3 is one of the points that makes GAMLSS modeling attractive, as it allows structured additive predictors (such as linear predictors, non-linear predictors, semi-parametric or non-parametric predictors such as splines or fractional polynomials of the independent variables and/or random effects terms) to model all the distribution parameters. In the absence of non-parametric additive terms in the model, we obtain the linear parametric GAMLSS model in its simplest form, which is the focus of this research. This structure allows for a fairer comparison with the Fama and French (2015) model, which models only the mean of the response. Thus, the systematic component to be used is described in Equation (4).

$$g_k(\theta_k) = \eta_k = X_k \beta_k, \tag{4}$$

and the covariates are replaced by the five factors of the Fama and French (2015) model. Thus, the average submodel will be

$$g_k(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5. \tag{5}$$

In Equation (5) the explanatory variables are the risk premiums. The parameters β_j, $j = 1, \ldots, 5$ determine the fixed effects of the risk premiums on the excess return for each portfolio analyzed. To ensure comparability with the Fama and French (2015) model, the scale and shape parameters were fixed in the distributions tested. It was found that different adjustment strategies can be considered, involving changes to the link functions and the inclusion of parametric or semi-parametric predictors for the scale and shape parameters, which can increase predictive power. On the other hand, this approach, although potentially more effective in terms of prediction, can face estimation problems, such as lack of convergence of the algorithm due to an insufficient amount of data or the loss of interpretability of the parameters, which is crucial in financial applications.

In the general case, the vector of parameters β_k and the random effects parameters γ_{jk} of the model in equation 3, for $j = 1, 2, \ldots, j_k$ and $k = 1, 2, 3, 4$, are estimated according to the GAMLSS framework by maximizing the penalized log-likelihood function l_p, defined by Equation (6)

$$l_p = l - \frac{1}{2} \sum_{k=1}^{p} \sum_{j=1}^{j_k} \lambda_{jk} \gamma'_{jk} G_{jk} \gamma_{jk}, \tag{6}$$

where $l = \sum_{i=1}^{n} log f\left(y_i \mid \theta^i\right)$ is the log-likelihood function of the model.

GAMLSS models offer significant advantages in financial market applications. Their distributional flexibility allows them to model unconventional distributions, such as those with heavy tails. In addition, these models have the ability to estimate a comprehensive set of parameters, including location, scale and shape. In this work, among the distributions tested are the t-family (TF), type 1 asymmetric t (ST1) and generalized t (GT), as they are robust versions of the normal distribution, behave with extreme values and were shown to more suitable for the data, assuming that the scale and shape parameters are fixed for all observations (constant scale and shape), based on the global deviation (GD), AIC and SBC selection criteria.

3.1.1 *T-family (TF) distribution*

The t-family distribution is defined as a distribution of three parameters: the mean (μ), scale (σ) and shape (τ). The density function of the t-family distribution is defined by Equation (7)

$$f\left(y|\mu,\sigma,\tau\right)=\frac{\Gamma\left(\frac{\tau+1}{2}\right)}{\sigma\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{\tau}{2}\right)\tau^{0,5}}\left[1+\frac{\left(y-\mu\right)^{2}}{\tau\sigma^{2}}\right]^{\frac{-(\tau+1)}{2}} \tag{7}$$

where $\mu\in R$, $\sigma>0$ and $\tau>0$ are the degrees of freedom. The $TF\left(\mu,\sigma,\tau\right)$distribution is symmetrical around $y=\mu$. Assuming $\mu=0$ and $\sigma=1$, we have the Student's t-distribution with τ degrees of freedom. The TF distribution also models leptokurtic data, i.e., data with a kurtosis greater than the kurtosis of the normal distribution (Rigby et al., 2019).

3.1.2 *Type 1 asymmetric t-distribution (ST1)*

The type 1 asymmetric t-distribution is a generalization of the normal asymmetric distribution and is a flexible and robust parametric distribution (Jones & Faddy, 2003). It allows for the adjustment of asymmetry and kurtosis (Álvarez & Gamero, 2012). The probability density function of ST1 can be expressed by

$$f\left(y|\mu,\sigma,\nu,\tau\right)=\begin{cases}\dfrac{c}{\sigma}\left[1+\dfrac{\nu^{2}z^{2}}{\tau}\right]^{\frac{-(\tau+1)}{2}} & sey<\mu\\[2ex]\dfrac{c}{\sigma}\left[1+\dfrac{z^{2}}{\nu^{2}\tau}\right]^{\frac{-(\tau+1)}{2}} & sey\geq\mu.\end{cases} \tag{8}$$

In Equation (8) the mean is $-\infty\langle\mu\rangle+\infty$, the scale is $\sigma>0$, $-\infty\langle\nu\rangle+\infty$, and $\tau>0$ for $-\infty\langle y\rangle+\infty$. Here, controls the tail of the distribution, i.e. large values of τ indicate heavier tails and therefore a greater probability of extreme events occurring, while ν controls the asymmetry.

3.1.3 *Generalized t-distribution (GT)*

This distribution is an extension of the Student's t-distribution and provides greater flexibility in data modeling, especially with respect to heavier tails (Verster & Waal, 2013). The probability density function of the generalized t-distribution is defined by Equation (9)

$$f\left(y|\mu,\sigma,\tau,\vartheta\right)=\vartheta\left\{2\sigma\tau^{\frac{1}{\vartheta}}\frac{\Gamma\left(\frac{1}{\vartheta}\right)\Gamma\left(\tau\right)}{\Gamma\left(\frac{1}{\vartheta}+\tau\right)}\left[1+\frac{z^{\vartheta}}{\tau}\right]^{\tau+\frac{1}{\vartheta}}\right\}^{-1} \tag{9}$$

where $z=\dfrac{\left(y-\mu\right)}{\sigma}$, in which -∞<y>+∞, the mean is $-\infty\langle y\rangle+\infty$, the dispersion is $\sigma>0$, $\tau>0$ and $\vartheta>0$. Here, τ controls the skewness and ϑ controls the tail of the distribution. The distribution $\left(\mu,\sigma,\tau,\vartheta\right)$is symmetric around $y=\mu$ and unimodal (Rigby et al., 2019).

As shown in Figure 1, the effect of the shape parameters on the distributions (GT, ST1 and TF) can be seen, reflected in the changes in kurtosis and asymmetry compared to the Gaussian pattern.

3.2 Sample

The data were taken from the Economatica database and consist of the historical monthly closing prices, adjusted for dividends and profits, of stocks

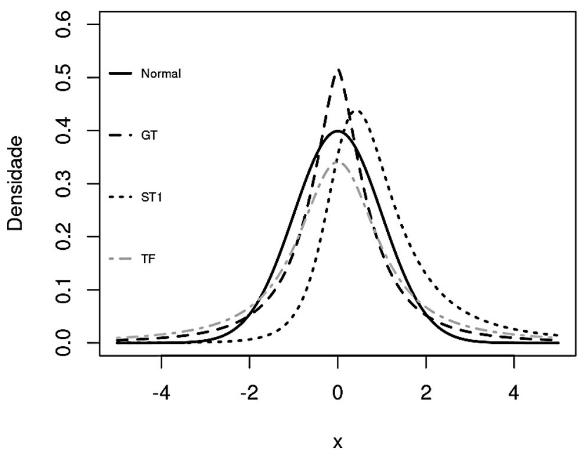

Figure 1. Normal, GT, ST1 and TF densities. Here, all distributions have $\mu=0$, $\sigma=1$ and ϑ =1.5

listed on the B3 (Brasil, Bolsa, Balcão). The analysis covers the post-1994 period, marked by the end of the hyperinflationary period in the Brazilian economy with the implementation of the Real Plan. Observations prior to this milestone were excluded from the sample, which ends in December 2018 (Supplementary Data 1 – Data).

4 Results

The results of the descriptive statistics of the analysis portfolios and risk premiums are available in the supplementary material and are not included in the article for reasons of concision (Supplementary Data 2 – R script).

4.1 HML: a redundant factor

Fama and French (2015) conclude that the five-factor model does not outperform the four-factor model, which excludes the risk premium for the book-to-market factor, denoted by HML in Equation (2). The authors' explanation is that the average return of the HML variable is captured by the variable's exposure to the risk premiums of the profitability and investment factors (RMW and CMA, respectively). For the American market, HML is redundant for the period between 1963 and 2013. Therefore, based on this empirical evidence, the redundancy assessment is carried out as proposed in Fama and French (2015). The assessment consists of estimating linear models via ordinary least squares (auxiliary regressions). In each model, one of the risk premiums is used as the dependent variable, while the other risk premiums are considered as regressors.

Table 3 shows the test results for the Brazilian data. In addition, the last column of this table shows the p-values of the Breusch-Pagan test (BP-test), which allows us to assess the presence of heteroskedasticity patterns in the

Table 3
Summary statistics using the four factors for different auxiliary regressions of the average returns of the fifth factor between January 1997 and December 2018

	Intercept	$R_m - R_f$	SMB	HML	RMW	CMA	Pseudo R^2	BP-test
$R_m - R_f$								
Coef	0.01		–0.01	–0.01	–0.02	0.00	0.002	
quasi-t	1.51		–0.54	–0.43	–0.56	0.05		
p-value	0.50		0.58	0.66	0.57	0.95		0.930
SMB								
Coef	–0.00	–0.03		–0.10	–0.03	–0.08	0.05	
quasi-t	–0.35	–0.51		–0.53	–0.20	–0.57		
p-value	0.72	0.61		0.59	0.83	0.56		0.00
HML								
Coef	0.01	–0.03	–0.14		–0.52	0.33	0.34	
quasi-t	0.66	–0.41	–0.54		–1.98	2.07		
p-value	0.50	0.67	0.58		0.04	0.03		0.00
RMW								
Coef	0.00	–0.06	–0.04	–0.44		0.19	0.23	
quasi-t	0.61	–0.65	–0.19	–2.21		0.8		
p-value	0.54	0.51	0.84	0.02		0.41		0.00
CMA								
Coef	–0.01	0.00	–0.19	0.51	0.35		0.21	
quasi-t	–0.47	0.06	–0.39	2.02	0.79			
p-value	0.63	0.94	0.69	0.04	0.42			0.00

residuals of the auxiliary regressions. The null hypothesis of homoskedasticity is rejected for the regressions with the variables SMB, HML, RMW and CMA as response variables. The only model that is homoskedastic is the one that includes the $R_m - R_f$ response variable. This justifies the use of the Horn et al. (1975) HC2 estimator to generate the quasi-t test statistics.

In the model where the dependent variable is $R_m - R_f$, the coefficient estimates in Table 3 show that the independent variables are not significant at the 5% level. In addition, Nagelkerke's (1991) pseudo-R^2 is close to zero, indicating that variations in the independent variables do not affect the dependent variable. The results indicate that the market premium is not a redundant factor. There is also insufficient evidence that the SMB variable is redundant. In the regression model with HML as the dependent variable, the coefficient estimates for RMW and CMA are -0.52% (quasi-t = -1.98 and p-value = 0.04) and 0.33% (quasi-t = 2.07 and p-value = 0.03), respectively, and the model has a pseudo-R^2 of 0.34. The regressions whose dependent variables are the risk premiums for profitability and investment (RMW and CMA) confirm the redundancy of the risk premium of the HML factor. Thus, it can be concluded that for the Brazilian financial market, the book-to-market risk premium, expressed by the HML variable, is partly explained by variations in the profitability and investment risk premiums, similar to the American financial market.

4.2 Modified five-factor model

A possible alternative to get around the redundancy problem would be to drop HML and use a four-factor model. Although this variable is captured by the exposure to other factors, Fama and French (2015, 2016) present arguments indicating that the book-to-market risk premium is important for assessing the behavior of asset prices. Thus, it is necessary to know the portfolios' exposure to the risk premiums for size, book-to-market, profitability and investment. The strategy adopted by Fama and French (2015) to overcome the redundancy of HML was also used in this study. The HML variable is replaced by the orthogonal HML variable (HMLO), which is the sum of the intercept and residual of the HML regression on $R_M - R_f$, SMB, RMW and CMA. Replacing HML with HMLO in Equation 2 yields the modified five-factor model (Equation (10)):

$$R_{i,t} - R_{f,t} = a_i + b_i\left(R_{m,t} - R_{f,t}\right) + s_i SMB_t + h_i HMLO_t + r_i RMW_t + c_i CMA_t + e_{i,t} \quad (10)$$

According to Fama and French (2015), both the intercept and the residual plot of Equation 10 are approximately the same as those presented in the five-factor regression (Equation 2), making the two regressions equivalent for analyzing the performance of the model.

4.3 GAMLSS models

In this section, by relaxing the normality assumption for the response variable, the GT, ST1 and TF distributions were used in the pricing models via GAMLSS, and the link function (*g*) used in the conditional mean models is the identity function. The scale parameters were assumed to be constant in order to make comparisons with the Fama and French (2015) model fairer. In addition, this approach is more parsimonious as it avoids including regressors for these parameters, thus maintaining an acceptable level of interpretability without compromising predictive power (Stasinopoulos et al., 2017). Furthermore, this decision is supported by the consideration of additional criteria that highlight the appropriateness and efficiency of the model, reinforcing the robustness of the methodological choice adopted (Regis, 2021). Further details on the selection of these distributions can be found in the supplementary material.

Table 4 shows the model estimates for portfolios formed by combining companies stratified by size and book-to-market. The intercepts of the models are generally close to zero. However, for the SN and BL portfolios, it is not possible to say that they are statistically zero, since the Z-statistics are -2.14 and -3.18, respectively (failing to reject the null hypothesis being tested).

The estimates of the market risk premium $(R_{m,t} - R_{f,t})$ are approximately equal to 1. The premiums associated with the size factor (SMB) are positive for small company portfolios and negative for big company portfolios, indicating a risk premium for investors in small company stocks. The book-to-market (HML) and investment (CMA) premiums are negative (positive) for companies with low (high) book-to-market, while the profitability premium (RMW) is positive (negative) for companies with low (high) book-to-market. The returns of portfolios of companies with low book-to-market are similar to those of companies with high profitability and aggressive levels of investment, while portfolios with high

Table 4
Regression estimates for the conditional mean (μ) in the best GAMLSS-fitted model for different combinations of size and book-to-market

$$g\left(\mu_{\left(R_t - R_{f,t}\right)}\right) = a + b\left(R_{m,t} - R_{f,t}\right) + sSMB_t + hHMLO_t + rRMW_t + cCMA_t$$

NE / MV →	Low (L)	Neutral (N)	High (H)	Low (L)	Neutral (N)	High (H)
MV		*a*			z(a)	
Small (S)	0.00	0.00*	0.00	1.34	–2.14	1.60
Big (B)	0.00*	0.00	0.00	–3.18	0.99	–0.08
		b			z(b)	
Small (S)	1.00*	0.93*	0.89*	35.76	30.00	38.95
Big (B)	0.94*	0.93*	0.96*	44.22	37.34	47.91
		s			z(s)	
Small (S)	1.24*	0.61*	0.81*	17.92	12.31	23.01
Big (B)	–0.15*	–0.09*	–0.08*	–4.57	–2.46	–4.36
		h			z(h)	
Small (S)	–0.89*	0.07	0.43*	–12.60	1.67	11.51
Big (B)	–0.28*	–0.02	0.45*	–9.22	–0.84	17.49
		r			z(r)	
Small (S)	0.32*	0.09*	–0.13*	6.11	2.64	–7.06
Big (B)	0.12*	–0.15*	–0.27*	5.09	–7.65	–10.48
		c			z(c)	
Small (S)	–0.17*	0.01	0.07*	–3.89	0.75	3.91
Big (B)	–0.17*	–0.07*	0.07*	–6.19	–4.67	6.29

Note: $|z|>1.64$ indicates significance at the 10% level and $|z|>1.96$ indicates significance at the 5% level. "*" indicates that the coefficient estimate is significant at the 5% level.

book-to-market have returns similar to those of companies with low profitability and conservative investment.

Table 5 shows the results for portfolios combining size and profitability. The inferences are similar to those in Table 4, with the intercept close to zero and the market risk approximately equal to 1. However, the "Big versus Robust" portfolio (acronym BR) shows an excess return similar to that of the stocks of profitable and aggressively invested companies, rejecting the null intercept hypothesis at a 5% significance level. The estimates for the size premium are similar, except for big companies with robust levels of profitability, which show a positive result.

The profitability premium estimates show that the premium is negative for companies with weak profitability, while the premium is positive for companies with robust profitability. As for the book-to-market and investment risk premiums, there is no clear pattern in the behavior of the estimates.

The results presented in Table 6 show a similar pattern for both the intercept and the market risk premium, which approach 0 and 1, respectively. However, when analyzing the "Big versus Neutral" (BN) portfolio, there is an excess return similar to that of stocks in profitable and aggressively invested companies, with the null hypothesis

Table 5
Regression estimates for the conditional mean (μ) in the best model fitted via GAMLSS for different combinations of size and profitability

$$g\left(\mu_{\left(R_t - R_{f,t}\right)}\right) = a + b\left(R_{m,t} - R_{f,t}\right) + sSMB_t + hHMLO_t + rRMW_t + cCMA_t$$

OP →	Weak (W)	Neutral (N)	Robust (R)	Weak (W)	Neutral (N)	Robust (R)
MV		*a*			z(a)	
Small (S)	0.00	0.00	0.00	–1.91	–0.81	1.51
Big (B)	0.00	0.00	0.00*	–0.09	0.00	–1.98
		b			z(b)	
Small (S)	0.90*	0.91*	0.95*	54.24	32.16	38.47
Big (B)	0.95*	0.85*	0.93*	36.57	27.97	45.92
		s			z(s)	
Small (S)	0.96*	0.70*	0.68*	45.77	10.65	8.86
Big (B)	–0.23*	–0.13*	0.06	–5.62	–8.02	1.16
		h			z(h)	
Small (S)	–0.14*	–0.01	0.14*	–5.72	–0.45	9.05
Big (B)	0.08*	0.04	–0.11*	2.22	0.96	–4.43
		r			z(r)	
Small (S)	–0.66*	0.00	0.34*	–26.83	0.01	30.89
Big (B)	–0.60*	–0.07*	0.20*	–17.82	–2.83	8.92
		c			z(c)	
Small (S)	0.03*	0.03	–0.02*	4.39	1.45	–2.21
Big (B)	0.07*	–0.32*	0.01	4.17	–19.53	0.67

Note: $|z|>1.64$ indicates significance at the 10% level and $|z|>1.96$ indicates significance at the 5% level. "*" indicates that the coefficient estimate is significant at the 5% level.

that the intercept is equal to zero being rejected at the 5% significance level. The estimates for the size factor risk premium are generally positive, except for large market value and conservative investment companies, where the estimate was not statistically significant. Companies with large market values had negative estimates for the book-to-market risk premium, while companies with small market values had negative risk premiums for profitability, indicating a negative premium for investors in stocks of companies with weak profitability. The investment risk premium estimates are positive for conservative investment firms and negative for aggressive investment firms, suggesting a positive excess return for investors in conservative firms.

The results of estimating the scale and shape parameters for the portfolios analyzed are available in the supplementary material. In addition, as in Moreira et al. (2021), the sample was divided into four sub-samples of five years each assessing the sensitivity of the estimates over time. The interpretations remained consistent, as detailed in the supplementary material.

4.4 Model performance evaluation

Table 7 presents the results for the performance metrics. Columns 2 to 6 correspond to the metrics used by Fama and French (2016). The standard linear regression models are estimated using ordinary least squares (OLS),

Table 6
Regression estimates for the conditional mean (μ) in the best model fitted via GAMLSS for different combinations of size and investment

$$g\left(\mu_{\left(R_t - R_{f,t}\right)}\right) = a + b\left(R_{m,t} - R_{f,t}\right) + sSMB_t + hHMLO_t + rRMW_t + cCMA_t$$

INV →	Conservative (C)	Neutral (N)	Aggressive (A)	Conservative (C)	Neutral (N)	Aggressive (A)
MV		*a*			z(a)	
Small (S)	0.00	0.00	0.00	–1.91	–1.92	–0.09
Big (B)	0.00	0.00*	0.00	1.10	–2.32	–1.35
		b			z(b)	
Small (S)	0.96*	0.91*	0.93*	38.77	38.19	33.31
Big (B)	0.89*	0.95*	0.92*	29.84	78.59	33.92
		s			z(s)	
Small (S)	0.89*	0.77*	0.84*	30.47	18.77	15.85
Big (B)	–0.03	0.05*	0.02	–0.59	2.45	1.69
		h			z(h)	
Small (S)	–0.12*	0.20*	0.06	–3.55	7.97	1.35
Big (B)	–0.07*	–0.07*	–0.32*	–2.77	–3.94	–12.21
		r			z(r)	
Small (S)	–0.19*	–0.08*	–0.05	–7.95	–4.33	–1.30
Big (B)	–0.05*	0.10*	–0.15*	–2.12	5.28	–9.23
		c			z(c)	
Small (S)	0.41*	–0.09*	–0.49*	23.55	–6.75	–11.72
Big (B)	0.34*	–0.08*	–0.60*	7.69	–5.87	–62.71

N **Note:** $|z|>1.64$ indicates significance at the 10% level and $|z|>1.96$ indicates significance at the 5% level. "*" indicates that the coefficient estimate is significant at the 5% level.

Table 7
Performance evaluation

Model	$A\lvert a_i\rvert$	$\frac{A\lvert a_i\rvert}{A\lvert \bar{r}_i\rvert}$	$\frac{A\lvert a_i^2\rvert}{A\lvert \bar{r}_i^2\rvert}$	$\frac{As^2 a_i}{A\lvert a_i^2\rvert}$	$A\left(R^2\right)$	A(MAE)
Standard Linear Regression	0.005	0.505	0.268	0.019	0.540	0.100
GAMLSS	0.005	0.502	0.271	0.047	0.810	0.090

and the GAMLSS models are estimated using likelihood maximization.

The evaluation $A|a_i|$ represents the average of the intercepts in absolute value. This metric is applied to both the standard linear regression and GAMLSS models. It can be seen that the values mentioned are close to zero, indicating that there is not a complete description of excess returns in all the models. This result is consistent with the studies of Fama and French (2015, 2016).

The results of the $\frac{A|a_i|}{A|\bar{r}_i|}$ metrics are 0.505 and 0.502 for the standard linear regression and GAMLSS modeling, respectively, indicating that in standard linear regression, the dispersion of the average unexplained excess returns is 50.5% greater than the dispersion of the average excess returns, values very similar to those obtained by GAMLSS (50.2%). However, when the analysis was performed in units of squared returns $\frac{A|a_i^2|}{A|\bar{r}_i^2|}$, 0.26 and 0.27 were obtained for standard linear regression and GAMLSS, respectively, i.e. the GAMLSS model showed a slightly higher value, with an average dispersion of unexplained excess returns that is 27% greater than the dispersion of average excess returns.

The metric $\frac{As^2 a_i}{A|a_i^2|}$ measures the extent to which the dispersion of the intercept is due to sampling error. In standard linear regression estimates, only 1.9% of the intercept dispersion is due to sampling error. On the other hand, in the GAMLSS models, the result is 4.7%, indicating that, with the best-fit distribution, there is a greater dispersion of the intercept due to sampling error, which is an improvement over standard linear regression.

When the analysis is performed based on the pseudo R^2 measurements, the following results are obtained: (1) for modeling using standard linear regression, the average value is 0.54, i.e., on average, the variations in the response variables are 54% explained by the risk premiums; (2) when considering the average of the pseudo R^2 of all the estimates obtained by the GAMLSS model, there is a pseudo-R^2 = 0.81, i.e., on average, the variation in the response variables is 81% explained by of the risk premiums; (3) for the linear regression, the average absolute mean error of all the portfolios is 0.10; (4) in the case of GAMLSS, the average mean error of the portfolios is 0.09. The MAE results show that the accuracy of the GAMLSS model is higher than that of the standard linear regression.

To verify the correct specification of the linear and GAMLSS models, the "worm plots" diagnostic tool is used according to Buuren and Fredriks (2001), using the residuals of the fitted model to reveal inconsistencies. The points indicate deviations of the residuals from the expected null value. A high (or low) variance results in positively (or negatively) skewed points. A U-shaped (or inverted) pattern of points indicates positive (or negative) asymmetry, while an S-shape with the left side up (or down), indicates high (or low) kurtosis. In a well-fitting model, most of the points lie within the semicircles. Figure 2 shows the worm plots for the models estimated by standard linear regression for the SL, SW and SC portfolios, highlighting the presence of high kurtosis with a negative slope in the residuals. Other similar plots are not included for concision.

Figure 3 shows the worm plots for the models estimated by GAMLSS modeling. Worm plots are not shown for all models, as they are similar to the three panels shown. In all cases, there is no evidence of location (mean) misspecification, as the residuals appear to be adequately distributed around zero. In addition, the plots of the residuals show no positive/negative slope, indicating that there is no misspecification of dispersion. Finally, the plots show no evidence of patterns of asymmetry or kurtosis

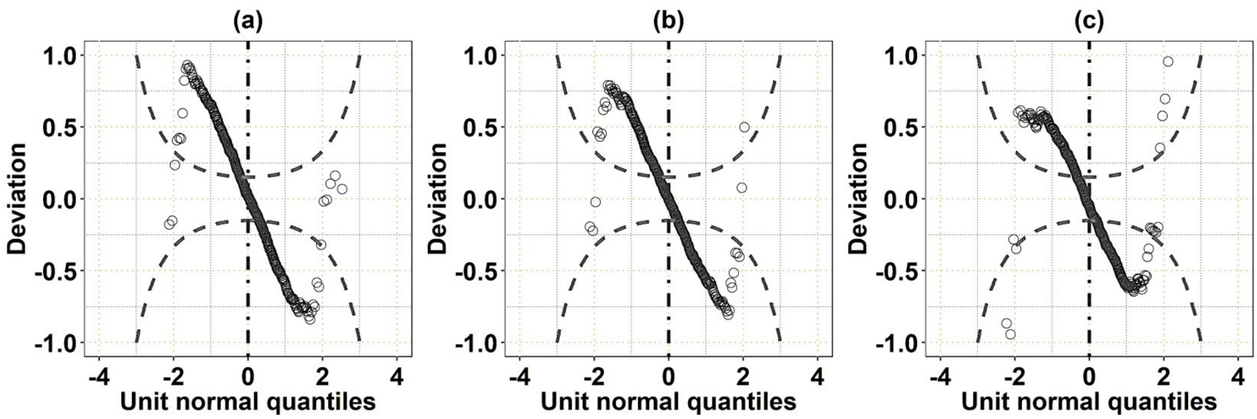

Figure 2. Worm plots where panel (a) is the SL portfolio, panel (b) is the SW portfolio, and panel (c) is the SC portfolio

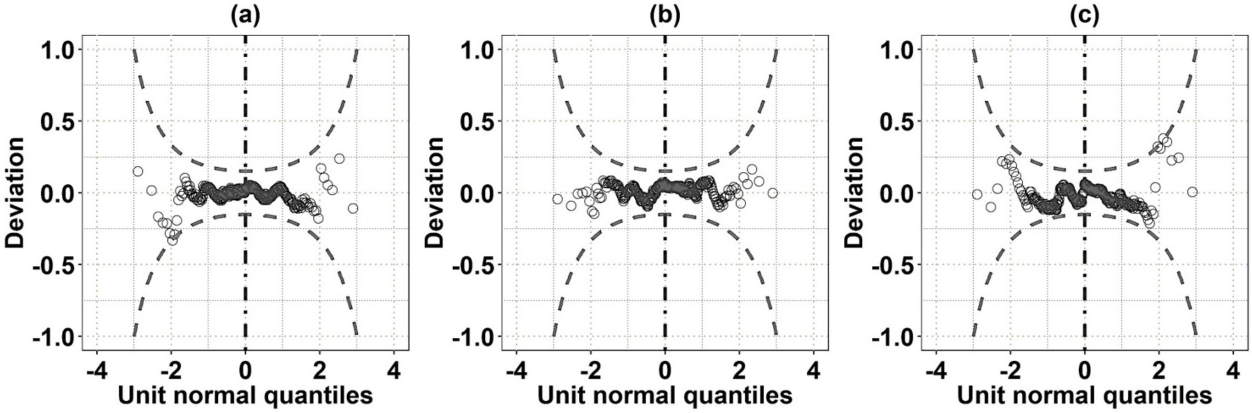

Figure 3. Worm plots where panel (a) is the SL portfolio, panel (b) is the SW portfolio, and panel (c) is the SC portfolio

that are not captured by the model. The worm plots shown in Figure 3 indicate that the GAMLSS models are well fit given the distributions fitted to the data.

The study highlights the superiority of GAMLSS over traditional models, confirming the conclusions of Regis et al. (2023) in the context of asset pricing, Florencio et al. (2012) in real estate pricing, and Matsumoto et al. (2022) in electricity pricing. These results highlight the innovative potential of this approach in asset pricing theory.

5 Conclusion

This article proposes the application of the GAMLSS econometric approach to asset pricing models, seeking a more appropriate choice in the absence of normality in the data. The research adopts a different approach to the study conducted by Regis et al. (2023), seeking a significant increase in the number of assets included in each portfolio.

In line with the observations of Fama and French (2015), the redundancy of the risk premium of the book-to-market factor is identified for the Brazilian data, captured by the exposure to the risk premiums of the profitability and investment factors. To overcome this problem, the HML variable was replaced by the orthogonal HML (HMLO), a transformation of the former.

When evaluating performance, it can be seen that all the estimated models provide incomplete descriptions of the excess returns of the portfolios. By using the appropriate distribution, the GAMLSS model improves the analysis of the intercept compared to models that use the normal distribution. In addition, the GAMLSS approach exhibits superior explanatory power, as evidenced by the pseudo coefficient of determination results. The mean absolute error values (an indicator of accuracy) show the superiority of the GAMLSS model. The worm plots show the specification error in the standard linear regression model, which is effectively corrected by the GAMLSS modeling.

GAMLSS modeling allows the incorporation of modeling for the mean and other parameters, by selecting a distribution that best fits the nature of the data, using scale and shape estimates to improve the fit. The contribution of this article lies in increasing the degree of diversification between portfolios, correcting the specification error of the models usually applied to Brazilian data, increasing the explanatory power of the application of pricing theory, and using distributions that take into account Brazilian stock behavior. This approach shows satisfactory results in the pricing of assets in the Brazilian financial market.

This work does not exhaust the full potential of the model, leaving ample room for econometric advances of this approach in empirical finance, for example, by estimating the GAMLSS model to verify the effect that risk premiums in Brazil have on the parameters of scale (dispersion) and shape (asymmetry and kurtosis). These parameters can provide information on how the behavior of financial market risk is influenced by market conditions and the intrinsic data of each company, and can be explored in future research. Other limitations of the research include the choice between numerous possibilities for the probability distributions associated with the response variables, as well as the computational execution time of the algorithms involved in the estimation process.

References

Álvarez, B. L., & Gamero, M. J. (2012). A note on bias reduction of maximum likelihood estimates for the scalar skew t distribution. *Journal of Statistical Planning and Inference*, *142*(2), 608-612. http://doi.org/10.1016/j.jspi.2011.08.012.

Buuren, S., & Fredriks, M. (2001). Worm plot: A simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*, *20*(8), 1259-1277. http://doi.org/10.1002/sim.746 PMid:11304741.

Carvalhal, A., & Mendes, B. V. (2003). Value-at-risk and extreme returns in Asian stock markets. *International Journal of Business*, *8*(1), 1-24. http://doi.org/10.2139/ssrn.420266.

Carvalho, G. A. D., Amaral, H. F., Pinheiro, J. L., & Correia, L. F. (2021). Precificação de anomalias através de modelos fatoriais: Um teste em mercados da América Latina. *Revista Contabilidade & Finanças*, *32*, 492-509. http://doi.org/10.1590/1808-057x202111640.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56. http://doi.org/10.1016/0304-405X(93)90023-5.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22. http://doi.org/10.1016/j.jfineco.2014.10.010.

Fama, E. F., & French, K. R. (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies*, *29*(1), 69-103. http://doi.org/10.1093/rfs/hhv043.

Florencio, L., Cribari-Neto, F., & Ospina, R. (2012). Real estate appraisal of land lots using GAMLSS models. *Chilean Journal of Statistics*, *3*(1). http://doi.org/10.48550/arXiv.1102.2015.

Hastie, T., & Tibshirani, R. (1995). Generalized additive models for medical research. *Statistical Methods in Medical Research*, *4*(3), 187-196. http://doi.org/10.1177/096228029500400302 PMid:8548102.

Horn, S. D., Horn, R. A., & Duncan, D. B. (1975). Estimating heteroscedastic variances in linear models. *Journal of the American Statistical Association*, *70*(350), 380-385. http://doi.org/10.1080/01621459.1975.10479877.

Jones, M. C., & Faddy, M. (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *65*(1), 159-174. http://doi.org/10.1111/1467-9868.00378.

Klein, N., Kneib, T., Lang, S., & Sohn, A. (2015). Bayesian structured additive distributional regression with an application to regional income inequality in Germany. *The Annals of Applied Statistics*, *9*(2), 1024-1052. http://doi.org/10.1214/15-AOAS823.

Koenker, R. (2005). *Quantile Regression, Volume Economic Society Monographs.* Cambridge University Press. http://doi.org/10.1017/CBO9780511754098.

Leite, A. L., Klotzle, M. C., Pinto, A. C. F., & da Silveira Barbedo, C. H. (2020). The Fama-French's five-factor model relation with interest rates and macro variables. *The North American Journal of Economics and Finance*, *53*, 101197. http://doi.org/10.1016/j.najef.2020.101197.

Matsumoto, T., Bunn, D., & Yamada, Y. (2022). Pricing electricity day-ahead cap futures with multifactor skew-t densities. *Quantitative Finance*, *22*(5), 835-860. http://doi.org/10.1080/14697688.2021.1984553.

Miller, M. H., & Scholes, M. (1972). Rates of return in relation to risk: A reexamination of some recent findings. *Studies in the Theory of Capital Markets*, *23*, 47-48.

Moreira, K. D. S., Penedo, A. S. T., Pereira, V. S., & Ambrozini, M. A. (2021). Crises e precificação de ativos no mercado de capitais brasileiro: Os cinco fatores de Fama & French. *Revista Gestão Organizacional*, *14*(2), 95-115. http://doi.org/10.22277/rgo.v14i2.4893.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691-692. http://doi.org/10.1093/biomet/78.3.691.

Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370-384. http://doi.org/10.2307/2344614.

Regis, R. O. (2021). *Precificação de ativos: uma análise dos cinco fatores de Fama e French em modelos GAMLSS* [Tese de doutorado]. Universidade Federal de Pernambuco.

Regis, R. O., Ospina, R., Bernardino, W., & Cribari-Neto, F. (2023). Asset pricing in the Brazilian financial market: Five-factor GAMLSS modeling. *Empirical Economics*, *64*(5), 2373-2409. http://doi.org/10.1007/s00181-022-02316-3.

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Applied Statistics*, *54*(3), 507-554. http://doi.org/10.1111/j.1467-9876.2005.00510.x.

Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., & De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R.* CRC press. http://doi.org/10.1201/9780429298547

Rocco, M. (2014). Extreme value theory in finance: A survey. *Journal of Economic Surveys*, *28*(1), 82-108. http://doi.org/10.1111/j.1467-6419.2012.00744.x.

Ross, S. (1976). The arbitrage pricing theory. *Journal of Economic Theory*, *13*(3), 341-360. http://doi.org/10.1016/0022-0531(76)90046-6.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425-442. http://doi.org/10.1111/j.1540-62611964tb02865.x.

Silva Jr, C. P. (2023). Uma comparação entre o modelo de três fatores de Fama e French e o modelo LCAPM no mercado acionário brasileiro. *Revista JRG de Estudos Acadêmicos*, *6*(13), 184-195. http://doi.org/10.5281/zenodo/.7957923.

Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., & De Bastiani, F. (2017). *Flexible regression and smoothing: using GAMLSS in R.* CRC Press.. http://doi.org/10.1201/b21973.

Verster, A., & Waal, D. J. (2013). The generalized t-distribution, a generalization to the positive tail of the t distribution. *South African Statistical Journal*, *47*(1), 71-82.

Vieira, M. D. V., Maia, V. M., Klotzle, M. C., & Pinto, A. C. F. (2017). Modelo de cinco fatores de risco: Precificando carteiras setoriais no mercado acionário brasileiro. *Revista Catarinense da Ciência Contábil*, *16*(48), 86-104. http://doi.org/10.16930/2237-7662/rccc.v16n48.2376.

SUPPLEMENTARY MATERIAL

Supplementary Data 1 – Data
Supplementary Data 2 – R Script
Supplementary material can be found online at https://doi.org/10.7910/DVN/AEGR7F

Financial support:
This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grant nos. 303192/2022-4 and 402519/2023-0, and Fundação de Amparo a Ciência e Tecnologia do Estado da Bahia (FAPESB), grant no. APP0021/20223 (R.O.).

Open science:
Renan Regis; Raydonal Ospina; Wilton Bernardino, 2024, "Asset Pricing: An Alternative Estimation for the Five-Factor Model," https://doi.org/10.7910/DVN/AEGR7F, Harvard Dataverse, V1

Conflicts of interest:
The authors have no conflicts of interest to declare.

Plagiarism analysis:
RBGN performs plagiarism analysis on all its articles at the time of submission and after approval of the manuscript using the iThenticate tool.

Authors:
1. Renan Oliveira Regis, PhD in Economics, Universidade Federal de Pernambuco, Recife, Brasil.
E-mail: renan.economia@hotmail.com
2. Raydonal Ospina, Post-Doctorate in Signal Processing, Universidade Federal de Alagoas, Maceió, Brasil.
E-mail: raydonal@de.ufpe.br
3. Wilton Bernardino, PhD in Economics, Universidade Federal de Pernambuco, Recife, Brasil.
E-mail: wilton.silva@ufpe.br

Authors' Contributions:
1st author: development of hypotheses and research questions (empirical studies); development of theoretical propositions (theoretical essays); theoretical foundation/literature review; data collection; statistical analysis; data analysis and interpretation; critical review of the manuscript; manuscript writing.
2nd author: definition of the research problem; development of hypotheses and research questions (empirical studies); development of theoretical propositions (theoretical essays); definition of methodological procedures; statistical analysis; analysis and interpretation of data; critical review of the manuscript; manuscript writing.
3rd author: definition of the research problem; development of hypotheses and research questions (empirical works); development of theoretical propositions (theoretical essays); statistical analysis; analysis and interpretation of data; critical review of the manuscript; manuscript writing.