JEAS, Vol. 31 No. 149 (2025) pp (73-91)
P-ISSN 2518-5764 - E-ISSN 2227-703X
DOI :https://doi.org/10.33095/9djpm450

Journal of Economics and Administrative Sciences

A Modified Fama-MacBeth Model based on the Single-Index Model

Mariam Jumaah Mousa *
Department Banking and Financial
Imam Alkadhim University College, Iraq

Munaf Yousif Hmood
Department of Statistics
College of Administration and Economics,
University of Baghdad, Iraq.

***Corresponding author**

Received:8/6/2025 **Accepted:3/8/2025** **Published: 1/10/2025**

Abstract:
The aim of this essay is to use a single-index model in developing and adjusting Fama-MacBeth. Penalized smoothing spline regression technique (SIMPLS) foresaw this adjustment. Two generalized cross-validation techniques, Generalized Cross Validation Grid (GGCV) and Generalized Cross Validation Fast (FGCV), anticipated the regular value of smoothing covered under this technique. Due to the two-steps nature of the Fama-MacBeth model, this estimation generated four estimates: SIMPLS(FGCV) - SIMPLS(FGCV), SIMPLS(FGCV) - SIM PLS(GGCV), SIMPLS(GGCV) - SIMPLS(FGCV), SIM PLS(GGCV) - SIM PLS(GGCV). Three-factor Fama-French model—market risk premium, size factor, value factor, and their implication for excess stock returns and portfolio returns—were estimated on the Iraqi Stock Exchange using the modified Fama-MacBeth. SIMPLS(FGCV)-GGCV performed best based on the findings. Results also revealed the statistical significance of the three factors of the Fama-French model, which enhanced the explanatory power of the model in terms of the performance of Iraq Stock Exchange
Keywords: Asset pricing model, The Fama-MacBeth Model Fama, French three factor model, Spline Smoothing, Single Index model, Generalized Cross Validation.

1. Introduction:

One of the simplest analysis tools to research variables that influence the financial markets and analyze financial returns is the Fama-MacBeth model. To make better investment decisions, forecast trends in the future, create investment choices, and carry out effective research of financial information to realize interactions between various economic components and how they influence markets as well as risk management; this model presents a risky tool. The cross-sectional regression is placed at the first stage of this two-stage process, and time series regression is then calculated from these estimators at the second stage. It is a major method in applied finance studies due to the possibility to separate the impact of single risks factor and anticipated to returns.

The Fama-MacBeth model has seen a plethora of estimation methods employed in recent literature. then Esteban González, M. V., & Orbe Mandaluniz, S. (2006) and Bailey, Calonacci, and Capitanius (2022) estimated the model initially through kernel weight regression and later through least squares to estimate the parameters. Ferreira, E., Gil-Baza, J., & Orbe, S. (2011). using smoothed least squares regression to estimate the parameters. While these methods are useful in the solution of some of the models' estimation issues, particularly where one has distributions of the model errors unknown or where one deals with time-varying data, they in fact require independence, normality of variables, and stationarity. These might not always be the case under extremely dynamic financial market conditions, particularly high volatility and uncertainty of emerging markets. According to this, the research issue stems from the necessity to create a more flexible and precise model to estimate the Fama-MacBeath model, particularly where it is applied in markets with specific characteristics such as the Iraqi market. The intricate and nonlinear interactions that could affect financial returns within such markets could not be overly sophisticated for traditional approaches to fully account for. Additionally, the means whereby the model can translate actual market performance convincingly and accurately may be limited by unyielding assumptions on the constancy of parameters or error distributions. Along this line, the research problem is the need to build a more flexible and effective system of estimating the Fama-MacBeth model, especially for use in markets with unique characteristics such as the Iraqi market.

Complexity and nonlinear relationships influencing financial returns in such settings might prove too complex for general tools to pin down conclusively.

Additionally, the ability of the model to fully describe actual market performance could be limited by dogmatic assumptions regarding parameter constancy and error structures. Therefore, by extending the Fama-MacBeth model with time-varying parameters and a link function, we want to improve its ability to account for market performance and offer the appropriate interpretation of Iraqi market determinants. In this adaptation, Grid Generalized Cross-Validation (GGCV) and Fast Generalized Cross-Validation (FGCV), two of the smoothing parameter methods utilized in the new model, will be compared. (comparison study between two smoothing parameter estimation procedures in the new model will be conducted).

Through applying both those approaches, the objective is to choose the most suitable and best parameter to enhance the accuracy of the model and reduce biases due to variations in financial information. This will enhance the validity of the outcome as well as the ability of the model to accurately and efficiently analyze market trends.

Previous studies have investigated the efficacy of the single index model using a variety of methods of estimation. Hmood, M. Y. (2015) estimated a link function in SIM using the Nadaraya-Watson estimator, Yu et al. (2002) estimated the single-indicator model using the penalized link (P-spline) estimation, and Hmood, M. Y., & Saleh, T. A. (2016). Compared the MAVE and RMAVE methods for estimating the single-indicator model without time-varying estimates, Ahmed, H. Y., & Hmood, M. Y. (2021) Comparison of Some Semi-parametric Methods in Partial Linear Single-Index by two stage procedure and MADE.

2. Literature Review and Hypothesis Development:

The FM approach has been used and modified in several studies to examine a variety of financial phenomena in different markets and periods of time.

For instance, Liu, Y. (2017) explored the impact of capital constraints on monetary policy transmission, finding that well-capitalized banks exhibit stronger transmission mechanisms, particularly for smaller banks. Abeland and Petersen (2018) used the FM approach to analyze compensation for exchange rate risk in international ETFs, concluding that US investors are compensated for such risk and that models combining equity and currency risk factors outperform solely equity-based models in explaining ETF excess returns.

The validity and explanatory power of various **asset pricing models** have been a significant area of research employing the FM methodology. Kim, B. W. (2018). tested the validity of asset pricing models in both US and Korean markets, affirming the pioneering role of Fama-French (1993, 2015) factor models due to their incorporation of asset size and value. The Fama-French three-factor model is a reasonably stable and relevant model, according to Lønø, B. E., & Svendsen, C. E. (2019) evaluation of various asset pricing models for the Norwegian stock market.

Nonetheless, certain markets have questioned the standard Capital Asset Pricing Model's (CAPM) relevance. Although there were indications of risk premiums linked to common skewness and downside beta, Markowski, L. (2020) concluded that the CAPM was not substantial in the Polish capital market El-Gebaly, A. A., & Essam El-Din. (2021). likewise concluded that the CAPM model was not valid for the Egyptian Stock Exchange, indicating that more sophisticated statistical techniques may be required for more precise estimations.

Beyond asset pricing, the FM model has been employed to investigate other market anomalies and information content. Yan, J. (2019) found that publicly available short-term flows, specifically the short volume ratio, are significant predictors of negative future stock returns. Chen, Y. M. (2019) utilized the FM method to estimate the information content of insider trading, exploring its relationship with investor behavior and distressed stocks. More recent applications include Nguyen, V. D. (2021), who examined the effect of product market competition on asset growth anomalies, and Kroon, E., & Karlsson, T. (2021), who attempted to replicate the retail trade imbalance anomaly in Swedish firms.

Improvements in the estimating methods for the Fama-MacBeth model are also demonstrated in the literature. While numerous research studies (e.g., Liu, Y. (2017); Apeland, M., & Pettersen, M. J. (2018).; Kim, B. W. (2018) Lønø, B. E., & Svendsen, C. E. (2019).; Yan, J. (2019); Chen, Y. M. (2019).; Markowski, L. (2020); El-Gebaly, A. A., & Essam El-Din. (2021).; Nguyen, V. D. (2021).; Kroon, E., & Karlsson, T. (2021); Khazanov, A. (2022); Persson, O., & Lindblom, S. (2024); Aghabeigi, M., & Ondes, T. (2024).) continue to use Ordinary Least Squares (OLS) as a common estimation method for the FM steps.

. Baillie, Calonaci, and Kapetanios (2022) and Esteban González, M. V., & Orbe Mandaluniz, S. (2006) both used kernel weight regression in the first phase and OLS in the second to estimate the model. Regression using smoothed least squares was used by Ferreira, E., Gil-Bazo, J., & Orbe, S. (2011). Though they frequently make assumptions about independence, standard distribution, and stationarity, which might not be true in extremely volatile or emerging markets, these alternative approaches seek to address issues like unknown error distributions or time-varying data.

The literature currently in publication emphasizes the significance of the Fama-MacBeth model while simultaneously pointing out the drawbacks of conventional estimation methods, particularly when used in markets with special features like the Iraqi market, which may show high volatility and nonlinear relationships. Since strict assumptions on error distribution or parameter stability may restrict the model's capacity to effectively interpret actual market performance, there is a clear need for more adaptable and precise frameworks to capture this complexity (Introduction).

2.1 Hypothesis Development:

The creation of a more reliable and adaptable framework is required due to the limitations of traditional Fama-MacBeth estimating methods for capturing the intricacies of emerging markets, particularly the Iraqi Stock Exchange. Using a Single Index Model (SIM) with time-varying parameters and a link function computed using penalized smoothing spline regression (SIMPLS), this study suggests modifying the Fama-MacBeth model.

The difficulties presented by possible non-linear relationships and non-stationary data in these markets are intended to be addressed by this method. A comparison analysis will be possible when selecting the best parameter for enhancing model accuracy and lowering biases through the use of two generalized cross-validation methods for smoothing parameter selection: Grid Generalized Cross-Validation (GGCV) and Fast Generalized Cross-Validation (FGCV).

Based on the proposed modification and the recognized need for improved estimation in volatile markets, the following hypotheses are formulated:

- **Hypothesis 1**: Compared to traditional OLS-based Fama-MacBeth estimations in the Iraqi Stock Exchange, the modified Fama-MacBeth model, which uses a Single Index Model (SIM) with time-varying parameters and penalized smoothing spline regression (SIMPLS), will offer a more reliable and accurate estimation of factor sensitivities and risk premiums.

- **Hypothesis 2:** The application of Generalized Cross-Validation methods (GGCV and FGCV) for smoothing parameter selection within the modified Fama-MacBeth SIMPLS framework will significantly improve the model's ability to explain the performance of the Iraqi Stock Exchange by better capturing non-linear relationships and reducing estimation biases.
- **Hypothesis 3:** Among the four potential estimation combinations resulting from the SIMPLS method with GGCV and FGCV for both steps of the Fama-MacBeth model (SIMPLS(FGCV)-SIMPLS(FGCV), SIMPLS(FGCV)-SIMPLS(GGCV), SIMPLS(GGCV)-SIMPLS(FGCV), SIMPLS(GGCV)-SIMPLS(GGCV)), one specific combination will empirically outperform the others in terms of explanatory power and statistical significance of the Fama-French three factors (market risk premium, size factor, and value factor) on excess returns for stocks and portfolios on the Iraqi Stock Exchange.

3. Research Methodology:

This section outlines the methodological approach employed in this study, focusing on the Fama-MacBeth regression model, the modified by Single Index Model with penalized smoothing spline regression (SIMPLS), and the estimation techniques for selecting smoothing parameters.

3.1 Fama MacBeth Regression Model (FM)

In 1973 the Nobel winner Eugene F. Fama and D. James MacBeth proposed a regression method for estimating cross-sectional regressions when one of the independent variables is a proxy (unobserved variable).

The Fama-MacBeth method involves two steps. The first step uses time-series regression to estimate coefficients (betas) for each factor. These betas then serve as independent variables in the subsequent cross-sectional regressions. For each dependent variable run the regression as shown in Equation (1). (Cochrane, J. (2009).; Lozano-Martín, 2009)

Suppose we have N dependent variables with k factors affecting the dependent variable.

$$Y_{i,t} = B_{0,i} + B_{1,i}X_{1,i,t} + B_{2,i}X_{2,i,t} + \cdots + B_{k,i}X_{k,i,t} + u_{i,t} \quad \ldots (1)$$

Where:

$Y_{i,t}$: is the dependent variable with $i = 1, 2, 3 \ldots N$; variables at time $t = 1, 2, 3, \ldots, T$.

$B_{k,i}$: is the kth coefficient of an independent variable $X_{k,i}$ at time t of dependent variable i.

$u_{i,t}$: is an independent and identical distributed error term with zero mean and variance $\sigma^2{}_{u_{i,t}}$.

In the second step, the estimated values of the coefficients in the first step (betas') will be regarded as an observation for the independent variable at each point in time to get an estimation of the coefficients in the second step (gammas') of FM regression model and then testing the significance of the model. (Baillie, et.at (2022)).

To explain the second step, we suppose the following model:

$$Y_{i,t} = \gamma_{0,t} + \gamma_{1,t}\hat{B}_{1,i,t} + \gamma_{2,t}\hat{B}_{2,i,t} + \cdots + \gamma_{K,t}\hat{B}_{K,i,t} + u_{i,t} \cdots (2)$$

Where: (Bai, J., & Zhou, G. (2015))

$\gamma_{0,t}$: is the intercept.

$\gamma_{1,t}$, $\gamma_{2,t}$,…, $\gamma_{k,t}$ are the coefficients of the independent variables (in period t) for the K factors

$\hat{B}_{K,i,t}$ independent variables estimate from step one.

Subsequently, a t-test statistic for these time-series estimates of gamma coefficients is computed to determine if the average gamma over time for each factor is significantly different from zero:

$$t_0 = \frac{\bar{\gamma}_{k,t}\ \sqrt{T}}{\sigma_{\gamma_{k,t}}} \cdots (3)$$

3.2 Modified the Fama-MacBeth Model by Single Index Model

In this section, we introduce a proposed approach for the Fama-MacBeth (FM) regression framework, leveraging the Single-Index Varying Coefficient Model estimated by penalized smoothing spline regression method (SIMPLS). This method is employed to estimate the model flexibly and to rigorously evaluate the statistical significance of its components over time.

A single index model is one of semiparametric models that are used as a compromise between restricted parametric models and a flexible nonparametric model. (Yu, Y., & Ruppert, D. (2002); Peng, H., & Huang, T. (2011).; Wu, J. (2015)).

To estimate time-varying parameters of equation (1) by SIM depending on the rolling regression method; SIM is formulated as:

$$Y_{it}(w) = g_{it}\left(X'_{itk}(w)B_{itk}(w)\right) + \epsilon_{it}\ (w)\ \ , i = 1,2..N\ , t = 1,2 \ldots T \cdots (4)$$

k=1,…, m is the number of independent variables.

N: is the number of dependent variables that refer to the number of observations in cross section model.

T: Time periods or total number of observations.

$g_{it}(.)$: is the unknown smooth link function of order w*1, the link function links the independent variables with the response variable through a single index.

$Y_{it}(w)$: is the dependent variable at time t for a window width w for variable i of order w*1.

$B_{itk}(w)$: is the index parameter of the dependent variable i at time t and independent variables k, representing the sensitivity of the factors influencing the dependent variable of order k×1 at times t and assume unknown $\eta_{it}(w) = X'_{itk}(w)B_{itk}(w)$ single index model ,$\|B\| = 1$ and its first component is greater than zero for the purpose of diagnosing the model

$X_{itk}(w)$: is the independent variables for origin i and window w at time t of order w*k

$\epsilon_{it}\ (w)$: is the random error at time t with order w*1.

To estimate the parameters of the second step of the Fama MacBeth methodology with cross section regression in (2), it is formulated as follows using the single index model as shown below:

$$Y_{it}(s) = g_{it}(\hat{\beta}'_{itk}(s)\gamma_{itk}(s)) + \epsilon_{it}\ (s), s = 1, 2\ldots T \cdots (5)$$

$Y_{it}(s)$:The dependent variable at time t for a cross sections s of degree N*1.

$\hat{\beta}'_{itk}(s)$: The sensitivity of the factors that are influencing the dependent variable with order N*k.

$\gamma_{itk}(s)$: is the index parameter of order k*1 and represents the sensitivity of the factors that influencing on the dependent variable and assume unknown used for interpretation once estimated.

3.3 Fama-MacBeth (FM) regression Estimation:

In this section we demonstrate (FM) estimation method using Spline Smoothing, where the goal is to estimate a smooth curve or function that best fits a set of observed data points.

Spline Smoothing is a powerful and effective tool in both time series regression and cross-sectional regression; it provides flexibility in adapting to irregular data and helps identify key patterns and trends. It can also be applied to estimate and smooth the underlying trend component in time series data, with the goal of removing noise or short-term fluctuations while preserving long-term patterns or trends. (Härdle, W. K. (2004).)

For the first step of Fama MacBeth, the Spline smoothing of the single index model can be considered as the residual sum of square represented by the difference between the real value of dependent variable and the link function in (5):

$$RSS = \sum_{j=1}^{w} \left(Y_{ijt}(w) - g_{ijt}\left(X'_{ijtk}(w) B_{ijtk}(w) \right) \right)^2 \cdots (6)$$

That is minimize the RSS but is merely interpolating the data, without exploiting any structure that might be present in the data. Spline smoothing solve this problem by adding a stabilizer or roughness penalty $\|g''_{ijt}\|_2^2$ that penalizes non-smoothness of $g_{ijt}(.)$.

$$\|g''_{ijt}\|_2^2 = \int \left(g''_{ijt}\left(X'_{ijtk}(w) B_{itk}(w) \right) = \hat{\eta}_{ijt}(w) \right)^2 d\hat{\eta}_{ijt}(w) \cdots (7)$$

Then the Spline smoothing is given as follows:

$$S_{\lambda_{it}}(g_{it}) = \sum_{j=1}^{w} \left(Y_{ijt}(w) - g_{ijt}\left(X'_{ijtk}(w) \hat{B}_{itk}(w) \right) \right)^2 + \lambda_{it} \left\| g''_{ijt} \right\|_2^2 \cdots (8)$$

Where the smoothing parameter λ is used to compromise between the goodness of fit and the continuity of the curve (penalty). (Maharani, M., & Saputro, D. R. S. (2021)).

The function estimator $g_{ijt}\left(X'_{ijtk}(w) \hat{B}_{itk}(w) \right)$ and the estimating parameter beta of first step time series regression can be determined by penalized least squares by minimizing (8).

For the second step of Fama MacBeth model, the Spline smoothing of the single index model can be consider as the residual sum of square represented by the difference between the real value of dependent variable and the link function in (6):

$$RSS = \sum_{j=1}^{s} \left(Y_{ijt}(s) - g_{ijt}\left(B'_{itjk}(s) \gamma_{itk}(s) \right) \right)^2 \cdots (9)$$

minimizes the RSS without utilizing any underlying structure that may be present. Spline smoothing addresses this issue by introducing a stabilizer that penalizes the non-smoothness $g_{ijt}(.)$

Where $\|g''_{ijt}\|_2^2$ is a roughness penalty, which is a measure of the smoothness of the curve in the data.

$$\|g''_{ijt}\|_2^2 = \int \left(g''_{ijt}\left(B'_{ijtk}(s) \hat{\gamma}_{itk}(s) \right) = \hat{\eta}_{ijt}(s) \right)^2 d\hat{\eta}_{ijt}(s) \cdots (10)$$

Then the spline smoothing is given as follows:

$$S_{\lambda_{it}}(g_{it}) = \sum_{j=1}^{s} \left(Y_{ijt}(s) - g_{ijt}\left(B'_{ijtk}(s) \hat{\gamma}_{itk}(s) \right) \right)^2 + \lambda_{it} \left\| g''_{ijt} \right\|_2^2 \cdots (11)$$

The function estimator $g_{ijt}\left(\hat{B}'_{itk}(s)\, \hat{\gamma}_{itk}(s) \right)$ and gamma parameters can be estimated by minimizing the penalized least squares of (11).

3.4 Asymptotic Analysis of the Penalized Least Squares (PLS):

The assumptions under we establish the consistency and find the rates of convergence of our estimators (Kuchibhotla, A. K., & Patra, R. K. (2020); Patra, R. K. (2016)). The link function is both non-constant and non-periodic, and it is differentiable almost everywhere. The first coordinate of parameter is positive, if λ goes to 0 at a rate faster than $n^{-1/4}$, but not faster than $n^{-2/5}$, then g(.) and B are consistent estimators of $g_0(.)$ and B_0, respectively.

$$\|\hat{g}_o\hat{\gamma} - g_o\gamma_o\| = O_p(\hat{\gamma}_s), \quad |\hat{\gamma} - \gamma_o| = O_p(\hat{\gamma}_s) \cdots (12)$$

For any function $g: R^d \to R, \|g\|^2 := \int g^2 dP\beta, (go\gamma)(\beta) = g(\gamma'\beta)$,and $P\beta$ denote the distribution of B if we choose $\hat{\lambda}_s = cs^{-\frac{2}{5}}, (c > 0)$ then both the prediction error and estimation error of $\hat{g}$ are of the optimal order i.e:

$$\|\hat{g}o\,\hat{\gamma} - go\gamma_o\| = O_p\left(s^{-2/5}\right) \cdots (13)$$

3.5 Selecting the Smoothing Parameter:

Selecting a smoothing parameter is critical for effective modeling and for balancing bias and variance, where a small smoothing parameter leads to under smooth, capturing noise in the data, while a large smoothing parameter leads to over smooth, potentially missing important features. There are several methods for selecting smoothing parameters, including. (Manaf, Y. H., & Mayassa, M. K. (2014); Liqa, A. M., & Kazim, S. H. (2017)):

3.5.1 Generalized Cross-Validation Method:

The choice of the smoothing parameter is based on minimizing the average-squared error of the fit of the model or based on adjustments of the error sum of squares, which yields an unbiased estimate of the prediction error.

$$GCV = \frac{\sum_{l=1}^{n}(Y_l - \hat{Y}_l)^2}{n(1 - \frac{tr(A_p)}{n})^2} \cdots (14)$$

Where:

Y_l : is the observed response value.

$\hat{Y}_l$: is the fitted response value.

n : is the number of observations.

A_p: represents the hat matrix or smoothing matrix and depends on x but not on y.

$\mathrm{tr}(A_p)$: is the trace of the matrix A_p which is equal to the degree of freedom.

To select the smoothing parameter λ in a single index model with time series regression by (Grid GCV (GGCV)) we use the fitted response of (4), where $\hat{B}_{it}(w)$ is estimated by minimizing (8) and the λ is equal to argmin of (14).

Otherwise, in cross sectional regression, the selection of the smoothing parameter λ in Single Index model by GGCV, depends on $\hat{B}_{it}(w)$ by minimizing (11) and λequal to argmin of (14).

ALGORITHM of GCV with Grid (GGCV)

This algorithm determines the optimal bandwidth for non-parametric regression. It is applied independently to each window in time series data and to each cross-section in cross-sectional data.

1. Initialize Grid Search: Define a dense grid H comprising a comprehensive range of possible bandwidth values h.
2. Iterative Model Estimation and GCV Calculation: For each h ∈ H:

a) Estimate the model using single index model (SIM)

b) Compute the Generalized Cross-Validation (GCV) function for the current h.

3. Optimal Bandwidth Determination: Identify the bandwidth h GCV that minimizes the calculated GCV function across all evaluated h values in H. $h_{GCV} = argmin_{h \in H}\, GCV\,(h)$

3.5.2 Bandwidth selection by using Fast method (GCVF):

In this method, the same formula is used (14) in all previous cases, except that the difference between them is the use of the ready-made package GCV in the program R which uses fit of model when estimating the parameters by ordinary least squares.

3.5.2 Fast Generalized Cross-Validation method (FGCV):

In this method, (15) is used as in all the previous cases, except that the difference between them is the use of the ready-made package GCV in the program R which uses fit of model when estimating the parameters by least squares.

3.6 Application:

In this section, we apply FM model with asset pricing model such as capital asset pricing, three factors Fama French model, multi-factor asset pricing models. (Baillie, R. T., Calonaci, F., & Kapetanios, G. (2022))

$$(R_i - R_f) = B_0 + B_1(F_{MARKET} - R_f) + B_2 F_{SMB} + B_3 F_{HML} + u_i \cdots (15)$$

Where:

R_i: is the actual return per share of the portfolio's stocks.

R_f: is the risk-free return is the government bond interest rate and resulting from the possibility of investing in assets where there is no possibility of their actual returns deviating from their expected returns.

$R_i - R_f$: is the stock risk premiums, that refers to the excess return or the additional return that exceeds the risk-free or base return of an investment. It refers also to the return generated by an investment after deducting the expected or base return. Excess returns are important because they show how an investment performs overall compared to the assumed or base rate of return that would have been achieved without taking a risk.

B_0: is a constant.

$B_i, i = 1,2,3$: is the slope coefficients for stock i and refers to the sensitivity of excess returns estimate of the i^th stock or stock portfolio to these factors (James W. Kolari · Wei Liu · Jianhua Z. Huang ,2021) or refers to the sensitivity of stock returns to changes in the corresponding factors is the risk for asset i associated with factor K, F_{MARKET} Market Return Index,$F_{MARKET} - R_f$ The excess return of the market portfolio represents the excess risk of the portfolio or represents the market premium over the riskless government bond rate (e.g., Treasury bill rate).

F_{SMB}: refers to the size of the company.

F_{HML}: is the risk factor for company growth that reflects to the ratio of book value to market value.

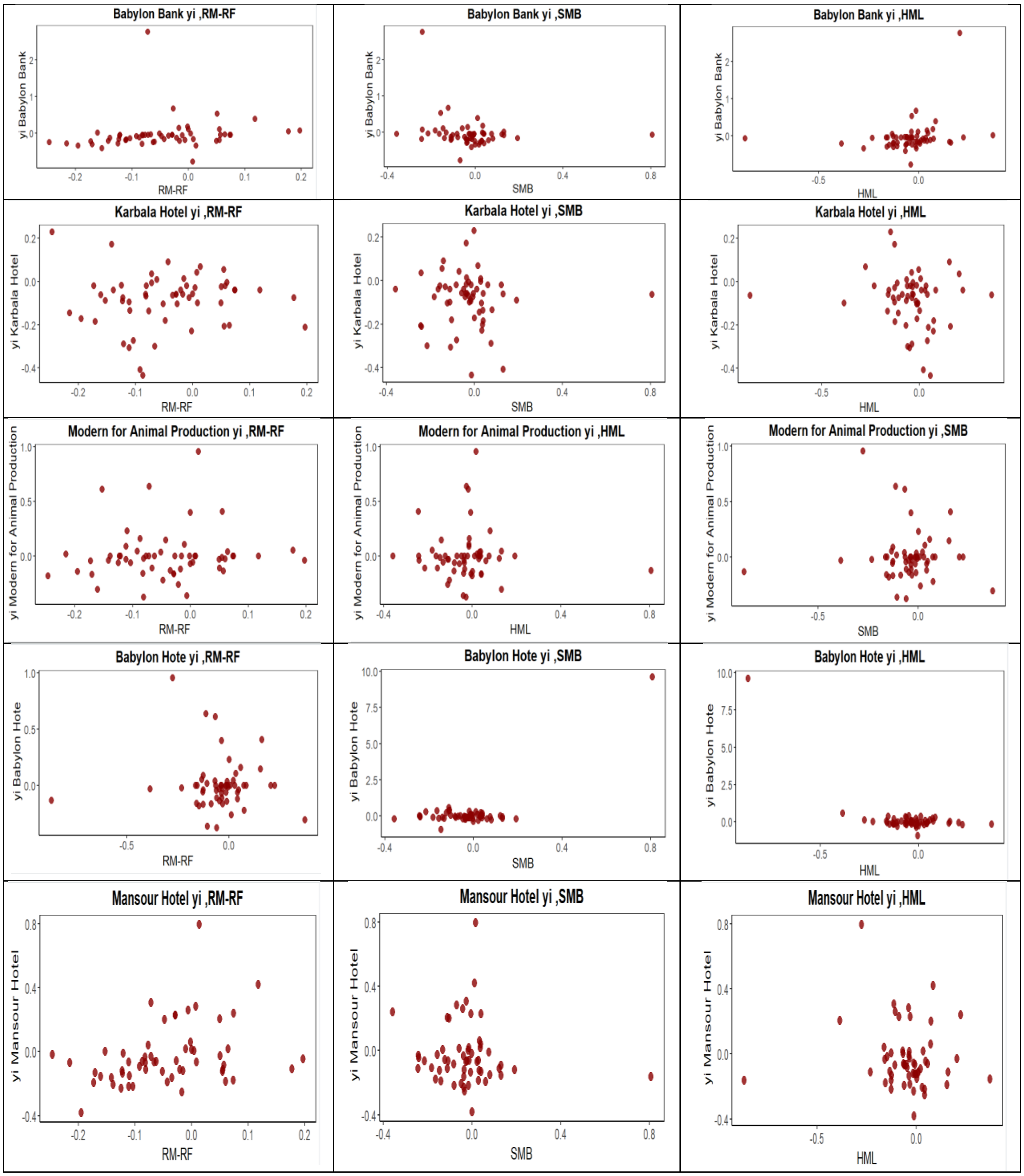
u_i: Unsystematic risk component of asset i, is the random error, which reflects the difference between the actual return on stock i in a given period and the return as predicted by the regression line.

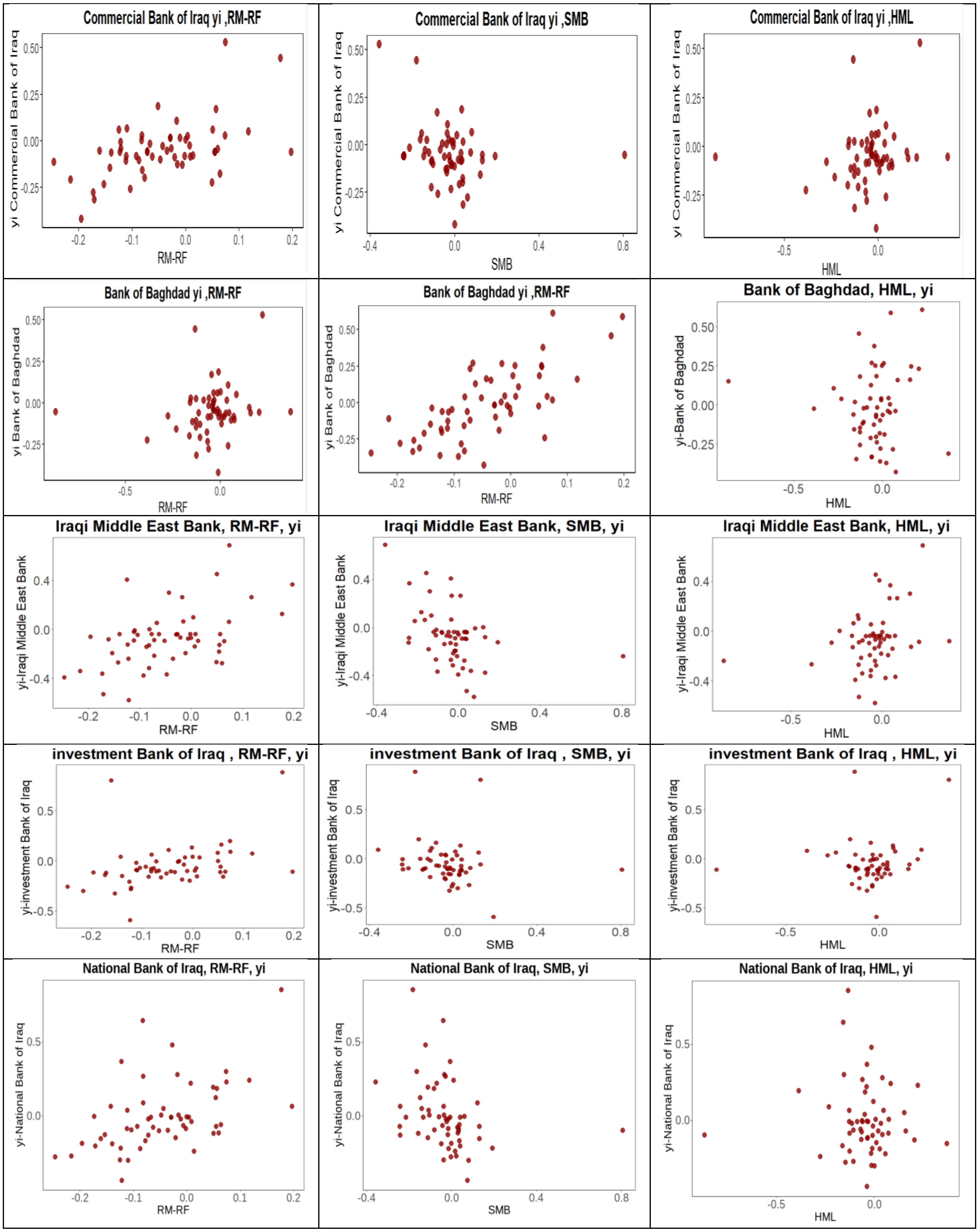
3.6.1 Data:

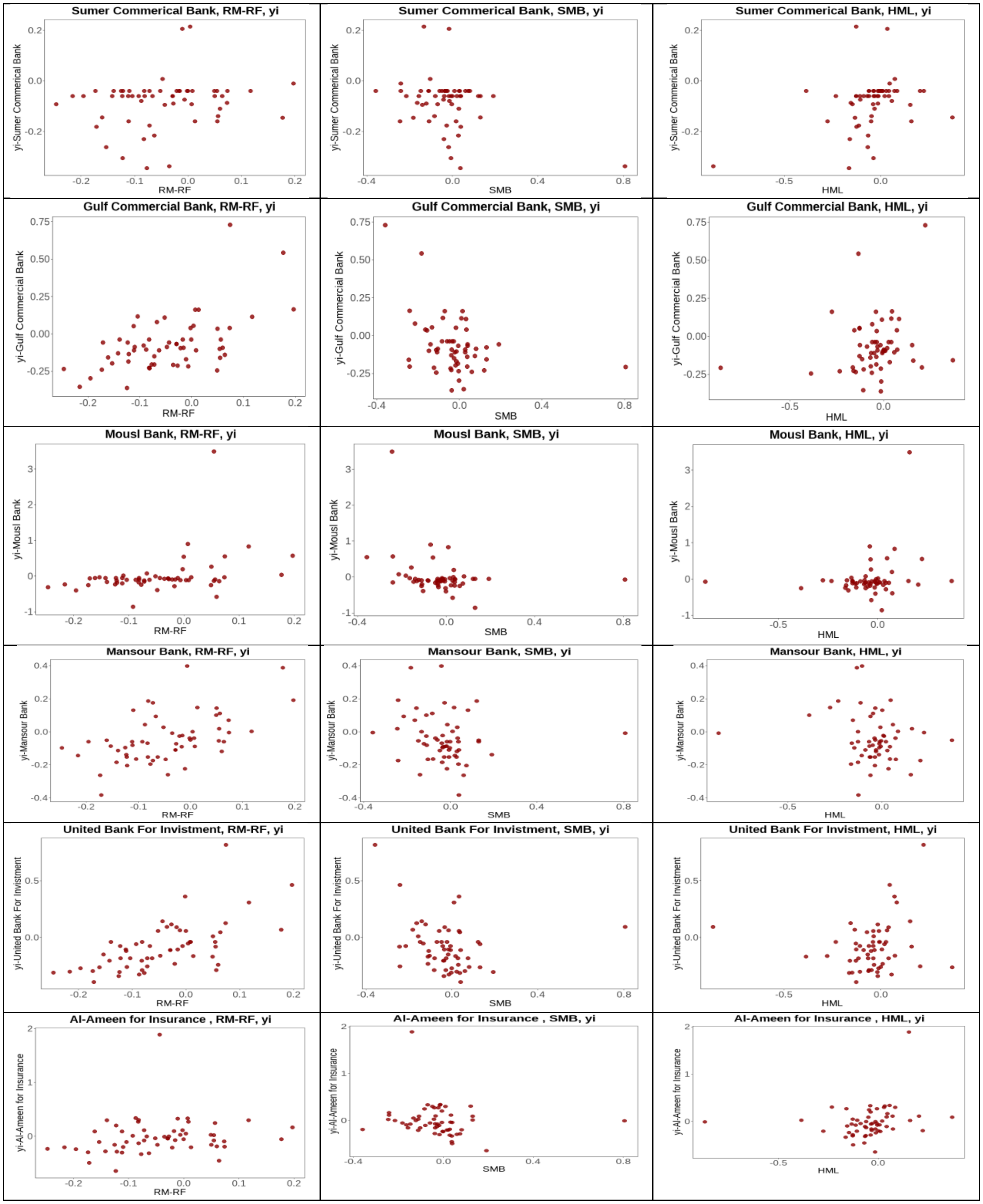
Regarding the research data, the data of the operating companies listed in the Iraq Stock Exchange were taken to give a description about the financial performance of the shares of some companies and try to achieve the goal of the study. Due to the availability of data in this market on FM variables. The sample consisted of 6 out of 8 sectors listed on the Iraq Stock Exchange, (banks, insurance, services, hotels, tourism, industry, and agriculture) because they are still operating and exist till now and have serial data that correspond to the time limits of the research, and we exclude the telecommunications and investment sectors because their data does not cover the research period. The research sample included (22) companies from these six sectors out of 103 operating companies. The data covered the time frame of the research which spanned for 14 years, from the first quarter of 2010 to the first quarter of 2024 (quarterly data) according to the data published in the company's directory of the Iraq Stock Exchange.

3.6.2 Test the Data:

The first step before estimating the model is to graphically examine the relationship between the explanatory (independent) variables and the dependent variable to ensure that the relationship is linear or non-linear. The R program was used to test the data by plotting the scatter plots. It was found that all relationships between the explanatory variables and the dependent variable are non-linear, as shown in the graphs below.




Commercial Bank of Iraq yi ,RM-RF
yi Commercial Bank of Iraq
RM-RF
Commercial Bank of Iraq yi ,SMB
yi Commercial Bank of Iraq
SMB
Commercial Bank of Iraq yi ,HML
yi Commercial Bank of Iraq
HML
Bank of Baghdad yi ,RM-RF
yi Bank of Baghdad
RM-RF
Bank of Baghdad yi ,RM-RF
yi Bank of Baghdad
RM-RF
Bank of Baghdad, HML, yi
yi-Bank of Baghdad
HML
Iraqi Middle East Bank, RM-RF, yi
yi-Iraqi Middle East Bank
RM-RF
Iraqi Middle East Bank, SMB, yi
yi-Iraqi Middle East Bank
SMB
Iraqi Middle East Bank, HML, yi
yi-Iraqi Middle East Bank
HML
investment Bank of Iraq , RM-RF, yi
yi-investment Bank of Iraq
RM-RF
investment Bank of Iraq , SMB, yi
yi-investment Bank of Iraq
SMB
investment Bank of Iraq , HML, yi
yi-investment Bank of Iraq
HML
National Bank of Iraq, RM-RF, yi
yi-National Bank of Iraq
RM-RF
National Bank of Iraq, SMB, yi
yi-National Bank of Iraq
SMB
National Bank of Iraq, HML, yi
yi-National Bank of Iraq
HML


Sumer Commerical Bank, RM-RF, yi
Sumer Commerical Bank, SMB, yi
Sumer Commerical Bank, HML, yi
yi-Sumer Commerical Bank
Gulf Commercial Bank, RM-RF, yi
Gulf Commercial Bank, SMB, yi
Gulf Commercial Bank, HML, yi
yi-Gulf Commercial Bank
Mousl Bank, RM-RF, yi
Mousl Bank, SMB, yi
Mousl Bank, HML, yi
yi-Mousl Bank
Mansour Bank, RM-RF, yi
Mansour Bank, SMB, yi
Mansour Bank, HML, yi
yi-Mansour Bank
United Bank For Invistment, RM-RF, yi
United Bank For Invistment, SMB, yi
United Bank For Invistment, HML, yi
yi-United Bank For Invistment
Al-Ameen for Insurance , RM-RF, yi
Al-Ameen for Insurance , SMB, yi
Al-Ameen for Insurance , HML, yi
yi-Al-Ameen for Insurance
RM-RF
SMB
HML

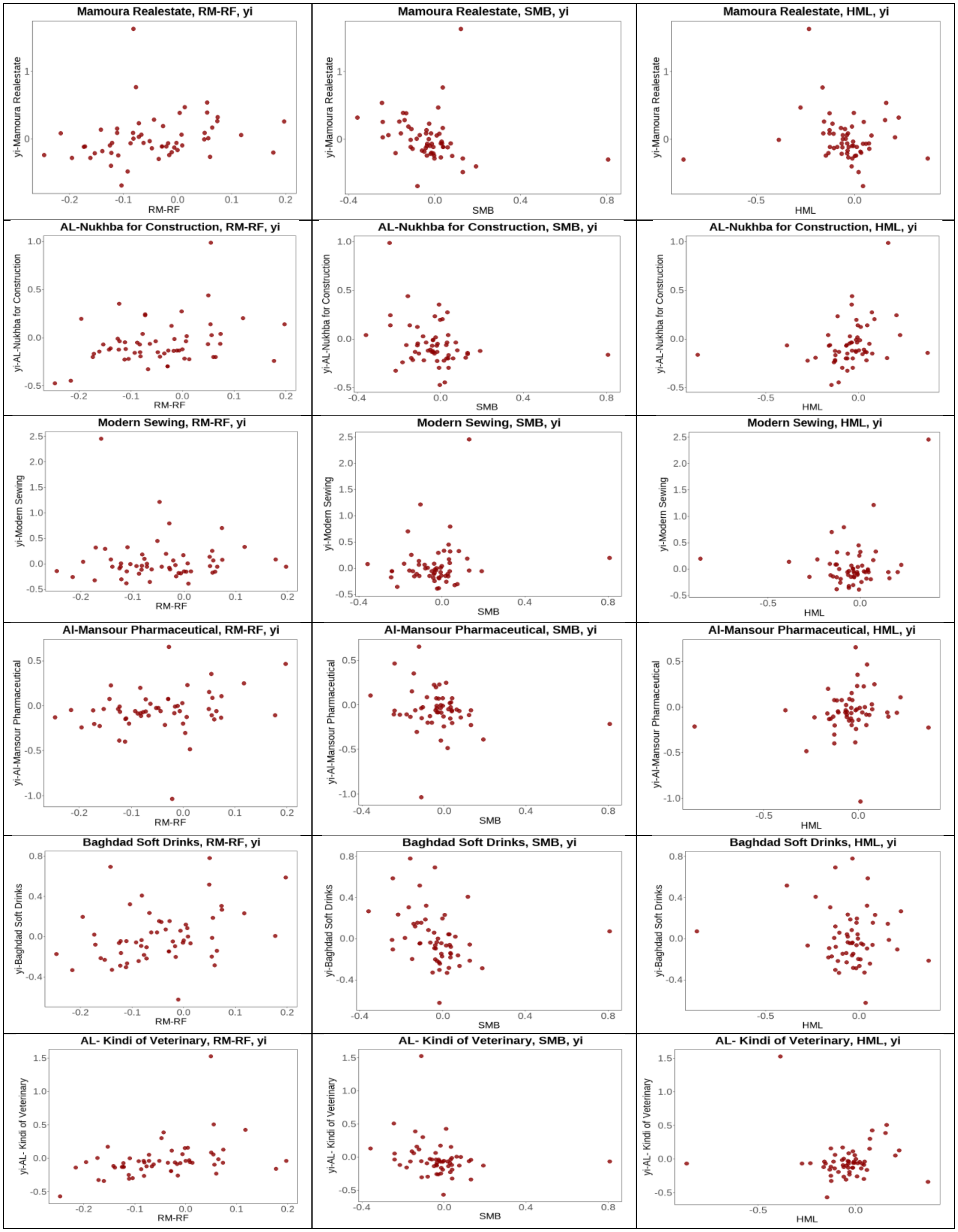


Figure (1): Relationship between the explanatory (independent) variables and the dependent variable of all assets

Source: researchers' preparation

4. Results:

Table (1,2,3) provides the results of model fit criteria for estimating the time varying beta at the first step of the FM based on a rolling window for five years with a single index model to get SIMPLS, and the time varying smoothing parameter according to the smoothing parameter that is selected by GGCV and FGCV.

We use VR1, which measures the goodness of fit as the ratio of the variance of estimated returns to the variance of actual returns. This serves as an indicator of how well the model explains the unconditional variance of returns. If the ratio is close to one, it indicates that the expected returns from the model closely match the actual returns from the statistical model.

We use also, VR2 to measure the estimation error that calculated as the ratio of the variance of the residuals to the variance of the actual returns. If this ratio is close to zero, it indicates that the model pricing errors are not significantly contributing to the predictable variation in the asset returns.

Table (1) below shows the VR1 criterion values for each asset, which compares the performance of the SIMPLS method using the smoothing parameter defined by GGCV and that defined by FGCV:

Table (1): VR1 criterion values for SIMPLS model

NAME OF ASSET	Statistic	SIM PLS (FGCV)	SIM PLS (GGCV)
Babylon Bank	VR1	0.21399	0.60461
Karbala Hotel	VR1	0.13245	0.5238
Modern for Animal Production	VR1	0.16698	0.20719
Babylon Hote	VR1	0.83651	0.96927
Mansour Hotel	VR1	0.22617	0.39948
Commercial Bank of Iraq	VR1	0.40507	0.73261
Bank of Baghdad	VR1	0.54308	0.6285
Iraqi Middle East Bank	VR1	0.61718	0.53764
investment Bank of Iraq	VR1	0.38446	0.66147
National Bank of Iraq	VR1	0.28707	0.32313
Sumer Commercial Bank	VR1	0.39945	0.57386
Gulf Commercial Bank	VR1	0.61201	0.53798
Mosul Bank	VR1	0.33519	0.22537
Mansour Bank	VR1	0.22203	0.25932
United Bank For Investment	VR1	0.64074	0.72572
Al-Ameen for Insurance	VR1	0.26518	0.42093
Mamoura Real estate	VR1	0.42062	0.46093
AL-Nukhba for Construction	VR1	0.3033	0.36568
Modern Sewing	VR1	0.35375	0.29316
Al-Mansour Pharmaceutical	VR1	0.11708	0.21838
Baghdad Soft Drinks	VR1	0.3839	0.38671
AL- Kindi of Veterinary	VR1	0.60718	0.87369

Source: researchers' preparation

Table (1) shows the VR1 values for the SIMPLS (GGCV) and SIMPLS (FGCV) methods for each asset. Analyzing these values, we note that many assets exhibit VR1 values closer to one using the SIMPLS (GGCV) method than SIMPLS (FGCV). For example, "Babylon Hotel" exhibits a VR1 value for SIMPLS (GGCV) of 0.96927, which is very close to one, indicating that the model explains a significant proportion of the variance in actual returns for this asset using GGCV. Assets such as "Commercial Bank of Iraq" and "Al-Kindi of Veterinary" also exhibit relatively good performance using SIMPLS (GGCV). In general, higher VR1 values for the SIMPLS (GGCV) method indicate a better ability to match expected returns with actual returns, supporting its superiority in explaining the unconditional variance of returns in the first step of the Fama-MacBeath model.

Table (2) presents the values of the VR2 criterion, which is an indicator of the estimation error, for each asset under both methods of selecting the smoothing factor:

Table (2): VR2 criterion values for SIMPLS model

NAME OF ASSET	Statistic	SIM PLS (FGCV)	SIM PLS (GGCV)
Babylon Bank		0.83448	0.16514
Karbala Hotel	**VR2**	0.81234	0.74154
Modern for Animal Production	**VR2**	0.84347	0.51355
Babylon Hote	**VR2**	0.16204	0.0135
Mansour Hotel	**VR2**	0.79883	0.33965
Commercial Bank of Iraq	**VR2**	0.48133	0.28557
Bank of Baghdad	**VR2**	0.40431	0.29968
Iraqi Middle East Bank	**VR2**	0.49843	0.39518
Investment Bank of Iraq	**VR2**	0.35321	0.21642
National Bank of Iraq	**VR2**	0.55798	0.40051
Sumer Commercial Bank	**VR2**	0.51987	0.37039
Gulf Commercial Bank	**VR2**	0.43087	0.31472
Mosul Bank	**VR2**	0.73615	0.54721
Mansour Bank	**VR2**	0.58963	0.44928
United Bank For Investment	**VR2**	0.42066	0.24386
Al-Ameen for Insurance	**VR2**	0.79242	0.79387
Mamoura Real estate	**VR2**	0.98088	0.52208
AL-Nukhba for Construction	**VR2**	0.52454	0.40894
Modern Sewing	**VR2**	0.81936	0.59734
Al-Mansour Pharmaceutical	**VR2**	0.7887	0.69334
Baghdad Soft Drinks	**VR2**	0.62615	0.50677
AL- Kindi of Veterinary	**VR2**	0.51971	0.10436

Source: researchers' preparation

Table (2) presents the values of the VR2 criterion, which reflects the estimation error, for both SIMPLS (GGCV) and SIMPLS (FGCV). Low VR2 values (close to zero) indicate that the model's pricing errors are small and do not significantly affects the expected variance of returns. Looking at the results, we find that the SIMPLS (GGCV) method exhibits VR2 values closer to zero for the majority of assets compared to SIMPLS (FGCV). For example, "Babylon Hotel" stands out with a very low VR2 value of 0.01350 for SIMPLS (GGCV), indicating high accuracy in its estimations. Similarly, "Al-Kindi of Veterinary" exhibits a low VR2 value (0.10436) with GGCV. These results reinforce the idea that using GGCV leads to lower estimation errors and thus improves the model's quality in determining the bootstrap parameter for the first step of the Fama-MacBeath model.

The following Table (3) shows the root mean square error (RMSE) values for each asset, providing an additional measure of estimation accuracy between the two methods:

Table (3): RMSE values for the SIMPLS model

NAME OF ASSET	RMSE SIM PLS (FGCV)	RMSE SIM PLS (GGCV)
Babylon Bank	0.2503	0.21215
Karbala Hotel	0.37559	0.08593
Modern for Animal Production	0.09947	0.1806
Babylon Hote	0.22635	0.18759
Mansour Hotel	0.32472	0.12673
Commercial Bank of Iraq	0.16809	0.09064
Bank of Baghdad	0.11465	0.12391
Iraqi Middle East Bank	0.14433	0.15136
investment Bank of Iraq	0.17013	0.09562
National Bank of Iraq	0.11985	0.17123
Sumer Commercial Bank	0.20211	0.05766
Gulf Commercial Bank	0.06759	0.11877
Mosul Bank	0.14036	0.22169
Mansour Bank	0.25535	0.09864
United Bank For Investment	0.11245	0.11284
Al-Ameen for Insurance	0.14707	0.17138
Mamoura Real estate	0.17046	0.16401
AL-Nukhba for Construction	0.22667	0.12364
Modern Sewing	0.14276	0.16956
Al-Mansour Pharmaceutical	0.20251	0.21075
Baghdad Soft Drinks	0.22539	0.17906
AL- Kindi of Veterinary	0.20008	0.09568

Source: researchers' preparation

Table (3) shows the RMSE values for each asset, which is a measure of accuracy, with lower values being preferable. Comparing the RMSE values between SIMPLS (GGCV) and SIMPLS (FGCV), we note that SIMPLS (GGCV) records lower RMSE values for most assets. For example, for Sumer Commercial Bank, the RMSE for SIMPLS (GGCV) was 0.05766, which is significantly lower than the RMSE for SIMPLS (FGCV) of 0.20211. This indicates that the estimates generated by SIMPLS (GGCV) are closer to the actual values, reflecting higher accuracy.

Overall, the RMSE results, along with the results from VR1 and VR2, support the conclusion that SIMPLS (GGCV) performs better in estimating the time-varying beta coefficient in the first step of the Fama-McBeath model, due to its higher accuracy and reduced estimation errors. Also, the overall average RMSE of SIMPLS(GGCV) (0.14316) is lower than the overall average RMSE of SIMPLS(FGCV) (0.18574), confirming the overall superiority of SIMPLS(GGCV).

In the cross-section regression of the individual dependent variable against beta time varying from first step of (FM) as independents variables.

The estimation of the modified Fama-MacBeth model yielded four distinct combinations, namely: SIMPLS (FGCV)-SIMPLS (FGCV), SIMPLS (FGCV)-SIMPLS (GGCV), SIMPLS (GGCV)-SIMPLS (FGCV), and SIMPLS (GGCV)-SIMPLS (GGCV)

Table (4): Estimation and Testing of Fama-French Model Parameters within the Modified Fama-MacBeth

Estimation Methods	**STATISTICS**	$\gamma_{0,t}$	$\gamma_{1,t}$	$\gamma_{2,t}$	$\gamma_{3,t}$
SIM PLS(FGCV)-SIM PLS(FGCV)	**Mean**	0.30761	0.39793	0.41151	**0.26100**
	S.E	0.06598	0.07024	0.05548	**0.03232**
	T-Stat	4.66202	5.66571	7.41696	**8.07501**
	P-Value	0.00004	0.00000	0.00000	**0.00000**
SIM PLS(FGCV)-SIM PLS(GGCV)	**Mean**	0.32449	0.41647	0.37465	**0.41591**
	S.E	0.02895	0.05064	0.06997	**0.04933**
	T-Stat	11.20631	8.22296	5.35400	**8.43030**
	P-Value	0.00000	0.00000	0.00000	**0.0000**
SIM PLS(GGCV)-SIM PLS(FGCV)	**Mean**	0.25317	0.26134	0.34438	**0.45151**
	S.E	0.02826	0.07509	0.05991	**0.06571**
	T-Stat	8.95845	3.48000	5.74804	**6.87092**
	P-Value	0.00000	0.00127	0.00000	**0.00000**
SIM PLS(GGCV)-SIM PLS(GGCV)	**Mean**	0.19136	0.26488	0.27701	**0.49904**
	S.E	0.03077	0.07854	0.07056	**0.05328**
	T-Stat	6.21914	3.37246	3.92600	**9.36658**
	P-Value	0.00000	0.00172	0.00035	**0.00000**

Source: researchers' preparation

From the Table (4) above, we note the significance of the parameters of the Fama-MacBeth model in the four methods. To determine the best method, we used Average Root Mean Squared Error (ARMSE).

Table (5): ARMSE of Fama MacBeth Regression

Methods	ARMSE
SIMPLS(FGCV)- SIMPLS(FGCV)	0.20433
SIM PLS(FGCV)- SIM PLS(GGCV)	0.12534
SIM PLS(GGCV)- SIM PLS(FGCV)	0.21315
SIM PLS(GGCV)- SIM PLS(GGCV)	0.12843

Source: researchers' preparation

ARMSE of the method SIMPLS(FGCV)-SIMPLS(GGCV), SIMPLS(GGCV)-SIMPLS(GGCV) in table (5) for cross-section regression are approximately and SIMPLS(FGCV)-SIMPLS(FGCV), SIMPLS(GGCV)-SIMPLS(FGCV) also approximate and the SIMPLS(FGCV)-SIMPLS(GGCV) is the smallest value present for the method. In any case, the presence of the grid in the cross-section regression had an important effect, while in the time series regression, no significant difference in performance was found between FGCV and GGCV.

After identifying the best method, represented by table (4) we now interpret the estimators for the Fama-MacBeth model.

$\bar{\gamma}_{0,t}$Its estimated value is 0.32449. This positive and significant value indicates the presence of an additional average return not fully explained by the three Fama-French factorsIn the Iraq Stock Exchange, this indicates the presence of residual market-specific risk factors that were not included in the model, or it may indicate some temporary market inefficiencies that allow for a fundamental return.

$\bar{\gamma}_{1,t}$ With an estimated value of 0.41647, this positive value confirms that investors in the Iraq Stock Exchange demand additional returns in exchange for bearing the overall systematic market risk. This is consistent with basic financial theory, which states that bearing market risk is rewarded with higher returns, highlighting the effectiveness of the market factor in asset pricing in the Iraqi market.

$\bar{\gamma}_{2,t}$ With an estimated value of 0.37465, this positive and significant value provides strong evidence that size plays a role in asset pricing on the Iraq Stock Exchange. This means that investors receive additional returns when investing in companies with small market capitalizations compared to companies with large market capitalizations.

$\bar{\gamma}_{3,t}$With an estimated value of 0.41591, its statistical significance and positive value indicate that investors in the Iraqi Stock Exchange demand additional returns when investing in value stocks compared to growth stocks.

5. Conclusions:

Based on the results, all proposed methods demonstrated that the three factors of the Fama-French model are statistically significant, enhancing the model's ability to explain performance in the Iraqi market.

The best method for estimating the Fama-MacBeath model, based on model fit criteria and accuracy measures, is the SIMPLS (FGCV)-SIMPLS (GGCV) method. Selecting the smoothing parameter using GGCV was more effective than FGCV, especially in the second step of the Fama-MacBeath estimation process. These findings provide valuable insights into advanced methods for modeling financial returns in emerging markets and the importance of selecting an appropriate smoothing parameter to enhance the accuracy of applied statistical models.

6. Recommendations:

The following recommendations are offered according to the study's findings:

1- Applying the Fama-MacBeath SIMPLS (FGCV)-SIMPLS (GGCV) model that has been modified: Given its higher reliability and precision than other tried-and-true techniques, this model is strongly advised for calculating financial returns and examining the elements that influence the Iraqi stock market.

2- The significance of selecting the smoothing parameter: Because the GGCV method has been shown to be successful in lowering estimation errors and improving the model's accuracy, researchers and analysts in emerging markets should give special consideration to the selection of the smoothing value.

3- Applying the model to other emerging markets: In order to confirm the generalizability of the Fama-MacBeath model and offer more comprehensive insights, this study recommends applying this updated approach to the model to other emerging financial markets with comparable features.

4- Future research could look into adding more risk factors (such momentum or profitability) to the SIMPLS framework in order to increase the model's explanatory power, even though the three Fama-French elements have proven to be significant.

5- Going deeper into the features of emerging markets: To develop more specialized and suitable asset pricing models, future research should keep examining the distinctive features of emerging markets and how they affect the relationship between risk and return factors.

Authors Declaration:

We hereby confirm that all the Tables in the manuscript are mine. Besides, the Figures and Images, which are not mine, have been permitted republication and attached to the manuscript.

- Ethical clearance: research was approved by The Local Ethical Committee in the University of Baghdad.

References:

Ahmed, H. Y., & Hmood, M. Y. (2021). Comparison of some semi-parametric methods in partial linear single-index model. *Journal of Economics and Administrative Sciences*, *27*(130), 170–184. https://doi.org/10.33095/jeas.v27i130.2207

Aghabeigi, M., & Ondes, T. (2024). Association between corporate governance and stock returns based on Fama and Macbeth two-stage regression (1973): A case study on companies in the Istanbul Stock Exchange. *International Journal of Nonlinear Analysis and Applications*, *16*(1), 123–133.

Apeland, M., & Pettersen, M. J. (2018). *Exchange rate risk compensation in international ETFs* (Master's thesis, Norwegian School of Economics).
https://brage.bibsys.no/xmlui/handle/11250/2559847

Bai, J., & Zhou, G. (2015). Fama-MacBeth two-pass regressions: Improving risk premia estimates. *Finance Research Letters*, *15*, 31–40.

Baillie, R. T., Calonaci, F., & Kapetanios, G. (2022). Hierarchical time-varying estimation of asset pricing models. *Journal of Risk and Financial Management*, *15*(1), 14. https://doi.org/10.3390/jrfm15010014

Chen, Y. M. (2019). *The information content of insider trading* (Doctoral dissertation, La Trobe University). https://doi.org/10.1016/j.frl.2017.12.007

Cochrane, J. (2009). *Asset Pricing* . Princeton University Press

El-Gebaly, A. A., & Essam El-Din. (2021). Testing the validity of capital asset pricing model (CAPM) in the Egyptian Stock Exchange. *Journal of Financial and Business Research*, *22*(3), 922–959.

Esteban González, M. V., & Orbe Mandaluniz, S. (2006). Nonparametric estimation betas in the market model. *Bilbao Research Papers*, *3*(1).
https://ideas.repec.org/p/ehu/biltok/5644.html

Ferreira, E., Gil-Bazo, J., & Orbe, S. (2011). *Conditional beta pricing models: A nonparametric approach*. SSRN. https://doi.org/10.2139/ssrn.1838375

Härdle, W. K., Werwatz, A., Müller, M., & Sperlich, S. (2004). *Nonparametric and semiparametric models*. Springer. https://doi.org/10.1007/978-3-642-17146-8

Hmood, M. Y. (2015). On single index semiparametric model. *Journal of Statistical Sciences*, *6*, 1–17.

Hmood, M. Y., & Saleh, T. A. (2016). Compared some of penalized methods in analysis the semi-parametric single index model with practical application. *Journal of Economics and Administrative Sciences*, *22*(90), 407–427. https://doi.org/10.33095/jeas.v22i91.490 .

Khazanov, A. (2022). *Nonlinear effects in international finance and macroeconomics.* (Doctoral dissertation, Boston College).

Kim, B. W. (2018). *Validity of the asset pricing models in applications to the U.S. and Korean markets.* (Master's thesis, California Los Angeles University).

Kroon, E., & Karlsson, T. (2021). *Replicating the retailers' trading imbalance anomaly: A quantitative study about excess return opportunities on Swedish small cap listed firms.* (Master's thesis, Uppsala University).

Kuchibhotla, A. K., & Patra, R. K. (2020). Efficient estimation in single index models through smoothing splines. *Bernoulli*, *26*(2), 1587–1618. https://doi.org/10.3150/19-BEJ1183

Liqa, A. M., & Kazim, S. H. (2017). Multi-linear in multiple nonparametric regression, detection and treatment using simulation. *Journal of Economics and Administrative Sciences*, *23*(101), 495–510. https://doi.org/10.33095/jeas.v23i101.190

Liu, Y. (2017). *Essays in financial crisis and capital regulations.* (Doctoral dissertation, University of California, Irvine).

Lønø, B. E., & Svendsen, C. E. (2019). *A comparison of asset pricing models in the Norwegian stock market.* (Master's thesis, BI Norwegian Business School).

Maharani, M., & Saputro, D. R. S. (2021). Generalized cross validation (GCV) in smoothing spline nonparametric regression models. *Journal of Physics: Conference Series*, *1808*(1), 012053. https://doi.org/10.1088/1742-6596/1808/1/012053

Manaf, Y. H., & Mayassa, M. K. (2014). Comparison of the estimations of a semi-parametric model using different bootstrapping methods. *Journal of Economics and Administrative Sciences*, *20*(75), 376–394. https://doi.org/10.33095/jeas.v20i75.595

Markowski, L. (2020). Further evidence on the validity of CAPM: The Warsaw Stock Exchange application. *Journal of Economics and Management*, *39*(1), 82–104. https://doi.org/10.22367/jem.2020.39.05

Nguyen, V. D. (2021). *The impact of competition on corporate decisions.* (Doctoral dissertation, The Australian National University).

Patra, R. K. (2016). *Semiparametric inference with shape constraints.* (Doctoral dissertation, Columbia University).

Peng, H., & Huang, T. (2011). Penalized least squares for single index models. *Journal of Statistical Planning and Inference*, *141*(4), 1362–1379. https://doi.org/10.1016/j.jspi.2010.10.003

Persson, O., & Lindblom, S. (2024). *Tangled up in metrics: A study on equity premiums in Europe.* (Master's thesis, Linnaeus University).

Wu, J. (2015). *Single-index regression models*. Indiana University.

Yan, J. (2019). *The informational content of short flow.* (Doctoral dissertation, Hong Kong University).

Yu, Y., & Ruppert, D. (2002). Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association*, *97*(460), 1042–1054. https://doi.org/10.1198/016214502388618861