INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104
Received: 11.11.2024 | **Accepted:** 20.03.2025 | **Published Online:** 30.03.2025

-RESEARCH ARTICLE-

CROSS-SECTIONAL REGRESSION WITH PROXIES: A SEMI-PARAMETRIC METHOD

Mariam Jumaah Mousa

Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq.
ORCID: https://orcid.org/0009-0006-6887-6297
Email: Mariam.Musa2201p@coadec.uobaghdad.edu.iq

Munaf Yousif Hmood

Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq.
ORCID: https://orcid.org/0000-0002-1134-9078
Email: munaf.yousif@coadec.uobaghdad.edu.iq

—Abstract—

This study investigates asset returns within the Iraq Stock Exchange by employing both the Fama-MacBeth regression model and the Fama-French three-factor model. The research involves the estimation of cross-sectional regressions wherein model parameters are subject to temporal variation, and the independent variables function as proxies. The dataset comprises information from the first quarter of 2010 to the first quarter of 2024, encompassing 22 publicly listed companies across six industrial sectors. The study explores methodological advancements through the application of the Single Index Model (SIM) and Kernel Weighted Regression (KWR) in both time series and cross-sectional analyses. The SIM outperformed the KWR approach in estimating time-varying beta coefficients, yielding a mean Root Mean Squared Error (RMSE) of 0.14316. Furthermore, the integrated KWR-SIM methodology achieved the lowest Adjusted Root Mean Squared Error (ARMSE) value of 0.08152 when modelling the association between risk factors and asset returns within the cross-sectional analytical framework. Statistical tests for significance produced heterogeneous responses of the returns on assets in the Iraqi financial market to the Fama-French posited economic variables. The estimated coefficients for the betas showed significant

Citation (APA): Mousa, M. J., Hmood, M. Y. (2024). Cross-Sectional Regression with Proxies: A Semi-Parametric Method. *International Journal of Economics and Finance Studies*, 17(01), 46-64. doi: 10.34109/ijefs.202517104

oscillations for all assets, confirming changes in economic conditions. The results add to our knowledge of the risk-reward relationship in the context of emerging markets and provide methodological insights into financial asset pricing. The evidence indicates that the KWR-SIM method has better capabilities for model fitting.

Keywords: Cross-Sectional Regression, Fama Macbeth Regression, Single Index Model, Kernel Weighted Regression, Proxy Variables, Time Varying Parameter.

INTRODUCTION

The analysis of cross-sectional regression represents a fundamental technique in finance and economics, employed to assess the relationship between variables across a group of units at a specific point in time (Pesaran, 2021; Wu et al., 2021). Nonetheless, researchers frequently encounter obstacles when the independent variables involved cannot be directly measured (Haslam et al., 2024). Qadri et al. (2024) address this issue by advocating the use of proxy variables, particularly where complex or latent variables are unobservable. The challenge intensifies when model parameters vary across time and seasons, rendering the analytical process significantly more intricate (Taheri et al., 2022; Toimil et al., 2020). Henley et al. (2020) suggest that addressing such dynamic frameworks necessitates the application of predictive statistical techniques capable of handling temporal variability. This study applies a cross-sectional regression framework that integrates time-variant model parameters and proxy independent variables. It seeks to establish reliable methodologies for interpreting complex interrelationships. To this end, the research adopts the Fama-MacBeth methodology, commonly referenced within financial literature, and supplements it with semi-parametric approaches, including kernel-weighted regression models as utilised by (González & Mandaluniz, 2006) and (Ferreira et al., 2011), who estimated Smooth Feasible Generalised Least Squares (SFGLS) and proposed a semi-parametric SIM.

The SIM, a standard instrument within multivariate regression, is utilized here to fit the sequential steps of Fama-MacBeth regression. Estimation of link function and SIM parameters is made with penalised smoothing splines (Aghabeigi & Ondes, 2024; Bryzgalova et al., 2024; Han et al., 2024). Smoothing splines are particularly effective in time series and cross-sectional regression analysis, offering flexibility in handling non-uniform data and the capacity to reveal underlying structural patterns (Härdle, 2004). The estimation of the nonparametric link function within the SIM structure is further explored. Yu and and Ruppert (2002) previously applied penalised spline (P-spline) estimation in partially single-index linear models and standard linear regressions. The Nadaraya-Watson estimator is employed to approximate the nonparametric function of the SIM, facilitating comparison between the MAVE and RMAVE methods used in estimating semi-parametric single-index models. This study develops four estimation configurations by combining KWR with the semi-parametric SIM: KWR-SIM, SIM-KWR, SIM-SIM, and KWR-KWR. The smoothing coefficients

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

and weighting strategies are optimised using the Generalised Cross-Validation (GCV) procedure. This research seeks to accomplish several core objectives:

- Firstly, it aims to develop robust estimation techniques capable of managing proxy variables and time-variant model coefficients within cross-sectional regression frameworks.
- Secondly, it intends to assess and compare the performance of the proposed methodologies, identifying the most effective approach in terms of estimation precision and model compatibility.
- Thirdly, the study applies these techniques to empirical data from the Iraq Stock Exchange, thereby investigating the real-world factors influencing this financial market.

This study provides scientific and practical insight into cross-sectional regression, particularly applicable to emergent and turbulent markets with sparse availability of direct data. Enhanced methodological precision enables researchers and financial analysts to better explain market behavior and make informed strategic decisions.

FAMA MACBETH REGRESSION MODEL (FM)

The Fama-MacBeth model establishes cross-sectional regression analysis as a fundamental econometric approach for employing independent variables as statistical proxies (Ferreira et al., 2011; Karolyi & Van Nieuwerburgh, 2020). Researchers apply this model to identify critical variables influencing their analyses, validate theoretical predictions (Green et al., 2016), and estimate parameters for models that incorporate time-varying dynamics (Škrinjarić, 2018). The model operates through a two-step procedure: initially, beta coefficients are derived via time-series regressions, followed by the estimation of gamma through cross-sectional regressions (Hodoshima et al., 2000; Jagannathan & Wang, 1998). While these procedures offer considerable analytical value, this study concentrates on enhancing the model's estimation techniques, recognising that the existing methods may present certain limitations.

THE CONSTRUCT AND ESTIMATING FAMA MACBETH TIME SERIES REGRESSION MODEL

The first stage of Fama-MacBeth regression entails a regression on a time-series to come up with proxy variables, which are then employed as the explanatory variables for cross-sectional regression, or for estimating time-varying beta (Fraser et al., 2004; Hur, 2024). This is typically achieved through rolling regression for each factor in the model, wherein the regression is conducted for all dependent variables, as demonstrated in the following Equation:

$$Y_{i,t}(w) = X_{it}(w)B_{it}(w) + \varepsilon_{it}(w) \ldots (1)$$

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

Where,

W: Window size or number of observations in window, with w from t-1 to t-w to estimate point t.
$Y_{i,t}(w)$: The dependent variable at time t for a window width w for variable i of degree w*1.
$B_{it}(w)$: The sensitivity of the factors influencing the dependent variable with a score of 2*1.
$X_{it}(w)$: The independent variables for origin i and window w at time t of degree w*2 for each factor.
$\varepsilon_{it}(w)$: The random error of the window (w) of dependent variables i at time t of degree w* 1with zero mean and constant variance.
Min $\varepsilon_{it}(w)' K_h(u) \varepsilon_{it}(w)$ … (2)
$K_{hw}(u)$: Kernel weights matrix of degree w*w with the Gaussian kernel function.
h: Bandwidth

$$\varepsilon_{it}(w)' K_h(u) \varepsilon_{it}(w) = \left(Y_{it}(w) - X_{it}(w) B_{it}(w)\right)' K_h(u) (Y_{it}(w) - X_{it}(w) B_{it}(w))$$
$$= Y_{it}(w)' K_h(u) Y_{it}(w) - Y_{it}(w)' K_h(u) X_{it}(w) B_{it}(w) - B_{it}(w)' X_{it}(w)' K_h(u) Y_{it}(w) + B_{it}(w)' X_{it}(w)' K_h(u) X_{it}(w) B_{it}(w)$$

By derivation of the last Equation with respect to $B_{it}(w)$ and setting it equal to zero, the research obtains an estimator $B_{it}(w)$.

$$\hat{B}_{it}(w) = (X_{it}(w)' K_h(u) X_{it}(w))^{-1} X_{it}(w)' K_h(u) Y_{it}(w) \dots (3)$$

To estimate the time-series regression with time-varying parameters of Equation (1) using SIM, the model is formulated as follows:

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

K=1...m: Number of independent variable.
N: Number of dependent variable and refer to number of observations in sectional.
T: Time periods or total number of views.
m: The number of factors influencing (independent variables) the dependent variable.

The function $g_{it}(.)$, a smooth but unknown link function, resembles the link function used in the general linear model, with the key distinction being that it is unspecified and must be estimated with a dimensionality of w*1. This link function serves a pivotal role in connecting the predictor variables to the response variable through a single index formulation. The variable $Y_{it}(w)$denotes the dependent variable for unit i at time t, corresponding to a window width w, with a dimensionality of w* 1. The parameter $B_{itk}(w)$ represents the index parameter associated with the dependent variable i, at time t, and factor k. It captures the sensitivity of the contributing factors to the

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

dependent variable, is of dimension k*1, and is assumed to be unknown prior to estimation. Once estimated, the single index model is given by $\eta_{it}(w) = X'_{itk}(w)B_{itk}(w)$, where ‖B‖=1, and the first component of B is constrained to be positive to facilitate model identifiability and diagnosis. The matrix $X_{itk}(w)$ corresponds to the independent variables for unit i, at time t, within the window w, having a dimensionality of w*k, where k denotes the number of factors. The term $\in_{it}(w)$ denotes the random error associated with unit i at time t, within window w, also of dimension w*1, and is assumed to have zero mean and finite variance.

The estimation of the link function and time-varying parameters in the Fama-MacBeth regression using SIM, through spline smoothing in the first step of the Fama-MacBeth model, can be considered as the residual sum of squares, represented by the difference between the actual value of the dependent variable and the link function in Equation (4). The model is defined as follows:

$$RSS = \sum_{j=1}^{w}\left(Y_{ijt}(w) - g_{ijt}\left(X'_{ijtk}(w)B_{ijtk}(w)\right)\right)^2 \dots (5)$$

Although this approach minimises the RSS, it merely interpolates the observed data without accounting for any underlying structural patterns that may exist. Spline smoothing addresses this limitation by introducing a regularisation term that penalises deviations from smoothness in the function $g_{ijt}(.)$, thereby stabilising the estimation process.

Here, $\left\|g''_{ijt}\right\|_2^2$ represents the roughness penalty, which serves as a quantitative measure of the smoothness of the curve used to model the data. This term penalises excessive curvature, thereby promoting a smoother functional form in the estimation of $g_{ijt}(.)$.

$$\left\|g''_{ijt}\right\|_2^2 = \int\left(g''_{ijt}\left(X'_{ijtk}(w)B_{itk}(w)\right) = \hat{\eta}_{ijt}(w)\right)^2 d\hat{\eta}_{ijt}(w) \ \dots (6)$$

The spline smoothing is then expressed as follows:

$$S_{\lambda_{it}}(g_{it}) = \sum_{j=1}^{w}\left(Y_{ijt}(w) - g_{ijt}\left(X'_{ijtk}(w)\hat{B}_{itk}(w)\right)\right)^2 + \lambda_{it}\ \left\|g''_{ijt}\right\|_2^2 \dots (7)$$

The parameter λ is employed to balance the trade-off between the goodness of fit to the data and the smoothness of the curve, serving as a smoothing parameter (Maharani & Saputro, 2021). The function estimator $g_{ijt}\left(X'_{ijtk}(w)\hat{B}_{itk}(w)\right)$ and the estimate parameter beta of first step time series regression can be determined by the used of penalized least square for minimizing Equation (7).

INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) **(pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

The second step involves performing a cross-sectional regression of the individual dependent variable against the proxy factor or beta as the independent variable at each point in time. This process aims to obtain a time series of gamma for each factor, with the average of these coefficients used to determine the parameters of the Fama-MacBeth regression, followed by testing their statistical significance (Ang & Kristensen, 2012; Liu & Huang, 2021). Using the provided dataset, perform the regression specified by the following Equation for each cross-section to estimate the coefficient $\gamma_{1,t=T}, \gamma_{2,t=T}$, $\gamma_{K,t=T}$, t=1,2…T time, as outlined in Equation (8). This approach allows for the estimation of time-specific parameter values across different cross-sectional units.

$$Y_{it}(s) = B_{itk}(s)\gamma_{tk}(s) + u_{itk}(s) \quad \ldots (8)$$

$Y_{it}(s)$:The dependent variable at time t for a cross sections s of degree N*1 .
N: Number of observations in sectional s.
$B_{itk}(s)$: Sensitivity of the factors influencing from time series regression the independent variable with a score of N*2.
$\gamma_{tk}(s)$: Amount of influence of the independent variable obtained in the first step on dependent variable.
$u_{itk}(s)$: Random error of the of degree N*1 with mean zero and finite variance.

Therefore, to estimate the parameters in Equation (8), the sum of squared residuals is minimised using all available observations.

Min $u_{itk}(s)'K_{hs}(U)u_{itk}(s)$ ... (9)
$K_{hs}(U)$:: Kernel weights matrix of degree N*N.
h: bandwidth

$$u_{itk}(s)'K_{hs}(U)u_{itk}(s) = (Y_{it}(s) - B_{itk}(s)\gamma_{tk}(s))'K_{hs}(U)(Y_{it}(s) - B_{itk}(s)\gamma_{tk}(s))$$
$$=R_{it}(s)'K_{hs}(U)Y_{it}(s) - Y_{it}(s)'K_{hs}(U)Y_{itk}(s)\gamma_{tk}(s) - \gamma_{tk}(s)'B_{itk}(s)'K_{hs}(U)Y_{it}(s) + \gamma_{tk}(s)'B_{itk}(s)'K_{hs}(U)B_{itk}(s)\gamma_{tk}(s)$$

By deriving the Equation and setting it equal to zero, the research obtains an estimator of $\gamma_{tk}(s)$.

$$\gamma_{tk}(s) = (B'_{itk}(s)\, K_{hs}(U)B_{itk}(s)\,)^{-1}B_{itk}(s)'K_{hs}(U)Y_{it}(s) \ldots (10)$$

To estimate the parameters of the second step in the Fama-MacBeth methodology, the cross-sectional regression in Equation (8) is formulated as follows, using the single index model as presented in the subsequent Equation.

$$Y_{it}(s) = g_{it}(B'_{itk}(s)\gamma_{itk}(s)) + \epsilon_{it}(s) \qquad s=1, 2...T \ldots (11)$$

The function $g_{it(.)}$ denotes an unknown smooth link function. The variable $Y_{it}(s)$, represents the dependent variable at time t for cross-section s, with a

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

dimensionality of N*1. The matrix ,$B'_{itk}(s)$ reflects the sensitivity of the factors influencing the dependent variable and is of dimension N* k. The vector $\gamma_{itk}(s)$ serves as the index parameter with dimensionality k*1, capturing the influence of explanatory variables on the dependent variable. This parameter is assumed to be unknown and is interpreted once estimated. The single index model is given by $\eta_s = B'_s\gamma_s$, and $\in_{it}$ (s)denotes the random error term, which is of dimension N* 1.

To estimate the link function and time-varying parameters of the Fama-MacBeth regression using SIM through spline smoothing in the second step of the Fama-MacBeth methodology, the residual sum of squares is considered, represented by the difference between the actual value of the dependent variable and the link function in Equation (11). The model is defined as follows:

$$RSS = \sum_{j=1}^{s}\left(Y_{ijt}(s) - g_{ijt}\left(B'_{itjk}(s)\gamma_{itk}(s)\right)\right)^2 \text{ ... (12)}$$

While this approach minimises the RSS, it merely interpolates the observed data without exploiting any underlying structural patterns that may be present. Spline smoothing addresses this issue by incorporating a stabilisation mechanism that penalises the non-smoothness of the function $g_{it}(.)$, thereby encouraging a more coherent and interpretable functional form.

Here, $\|g''_{ijt}\|_2^2$ represents the roughness penalty, which quantifies the degree of smoothness in the curve that fits the data. A lower value indicates a smoother curve, thereby helping to prevent overfitting by penalising excessive curvature in the estimated function.

$$\|g''_{ijt}\|_2^2 = \int\left(g''_{ijt}\left(B'_{ijtk}(s)\hat{\gamma}_{itk}(s)\right) = \hat{\eta}_{ijt}(s)\right)^2 d\hat{\eta}_{ijt}(s) \text{ ...(13)}$$

The spline smoothing is then expressed as follows:

$$S_{\lambda_{it}}(g_{it}) = \sum_{j=1}^{s}\left(Y_{ijt}(s) - g_{ijt}\left(B'_{ijtk}(s)\hat{\gamma}_{itk}(s)\right)\right)^2 + \lambda_{it}\ \left\|g''_{ijt}\right\|_2^2 \text{ ... (14)}$$

Where, λ is the smoothing parameter.

The function estimator $g_{ijt}\left(\hat{B}'_{itk}(s)\hat{\gamma}_{itk}(s)\right)$ and the parameters γ can be estimated using penalised least squares by minimising Equation (14). This approach balances model fit with curve smoothness through a penalty term. The *slimmest* package in R (Patra, 2016) offers an efficient algorithm for computing this estimator, particularly in scenarios involving moderately high-dimensional data (e.g., d = 100).

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

GENERALIZED CROSS-VALIDATION

The bandwidth is selected based on minimising the residual sum of squares, which provides an unbiased estimate of the prediction error, as outlined by Aydına et al. (2016) and Köhler et al. (2014).

$$GCV = \frac{\sum_{l=1}^{n}(Y_l - \hat{Y}_l)^2}{n(1-\frac{tr(A_p)}{n})^2} \quad \ldots (15)$$

Where,

Y_l: The observed response value.

$\hat{Y}_l$: The fitted response value obtained for the observed predictor vector.

n= The number of observations of the variable Y_l.

A_p: The hat matrix (also known as the smoother matrix), which depends on the predictor vector x, but not on the response variable y.

A_p: The trace of the matrix A_p, which is equal to the degrees of freedom.

To select the bandwidth h in KWR for time series regression, the study uses the fitted response from Equation (1), with $\hat{B}_{it}(w)$ estimated via Equation (3), and h chosen to minimise Equation (15). For cross-sectional regression, h is based on the fitted response from Equation (8), using γ from Equation (10), and similarly selected by minimising Equation (15). For the smoothing parameter γ in the single index model with time series regression, the fitted response from Equation (4) is used, with $\hat{B}_{it}(w)$ obtained by minimising Equation (7); λ is then selected by minimising Equation (15). In the cross-sectional case, λ is determined using the fitted response from Equation (11), where $\hat{B}_{it}(w)$ is estimated by minimising Equation (14), with λ also chosen as the minimiser of Equation (15).

APPLICATION OF FAMA MACBETH

The present study evaluates asset returns on the Iraq Stock Exchange using the Fama-MacBeth model and the Fama-French three-factor model. The primary objective is to assess asset risk by considering the Market Return Index, book-to-market value, and company size. The model is applied to 22 publicly listed companies from six of the eight operating sectors, utilising quarterly data from the first quarter of 2010 to the first quarter of 2024. The Market Return Index, one of the independent variables in the Fama-French model, measures systematic risk that cannot be mitigated through diversification. It reflects the extent to which the returns of a specific asset fluctuate in response to changes in the market return. This index serves as a gauge of market risk for the stock, with an assumed direct relationship between the Beta 1 coefficient and the required return. It indicates the degree of risk that the stock contributes to the market portfolio. The company size factor F_{SMB} represents a crucial determinant of financial management within the company. The book-to-market value ratio (FHML), another key risk factor in the Fama-French model, represents a company's growth potential, determined by the ratio of book value to market value (BE/ME). A high BE/ME ratio signifies a consistent decline in earnings per share, suggesting that stocks with a high BE/ME are less profitable than those with a low BE/ME. A low BE/ME ratio is typical of companies with high average returns on capital (growth stocks), while a high BE/ME ratio is common among companies facing relative distress (value stocks) (Fama & French, 1995).

The first practical step in this study involves estimating time-varying beta values (proxy variables) using two distinct approaches: the SIM and KWR. The research employs two metrics to assess the model's goodness of fit for each approach: the VR1 criterion, which confirms the significance of the goodness of fit, and the VR2 criterion, which gauges estimation error. The accuracy of the estimates is further assessed by calculating the RMSE. As presented in Table 1, the VR1, VR2, and RMSE values are extracted for both the KWR and SIM methods. According to the VR1 criterion, values closer to one are considered better, while values approaching zero for VR2 and RMSE indicate greater estimation accuracy. As highlighted by (Ferson & Harvey, 1991), the values in Table 1 reflect the goodness of fit for both methods.

Table (1) shows that when estimating the SIM compared to the KWR, the VR1 values are closer to one, and VR2 values near zero for most assets. The RMSE for each asset was calculated by comparing estimated and actual values (Hodson, 2022). The results indicate that the average RMSE for the SIM (0.14316) is smaller than for KWR. Figure (1) illustrates the time-varying beta estimates for listed companies using the SIM method. Based on the evaluation in Table (1), SIM was the most successful approach. The graph displays the time-varying beta values for all assets, showing three risk factors over a five-year window (Q1 2010 to Q4 2014), with 38 beta values for each variable. The black, red, and blue lines represent b1, b2, and b3, respectively, for the period from Q4 2014 to Q1 2024.

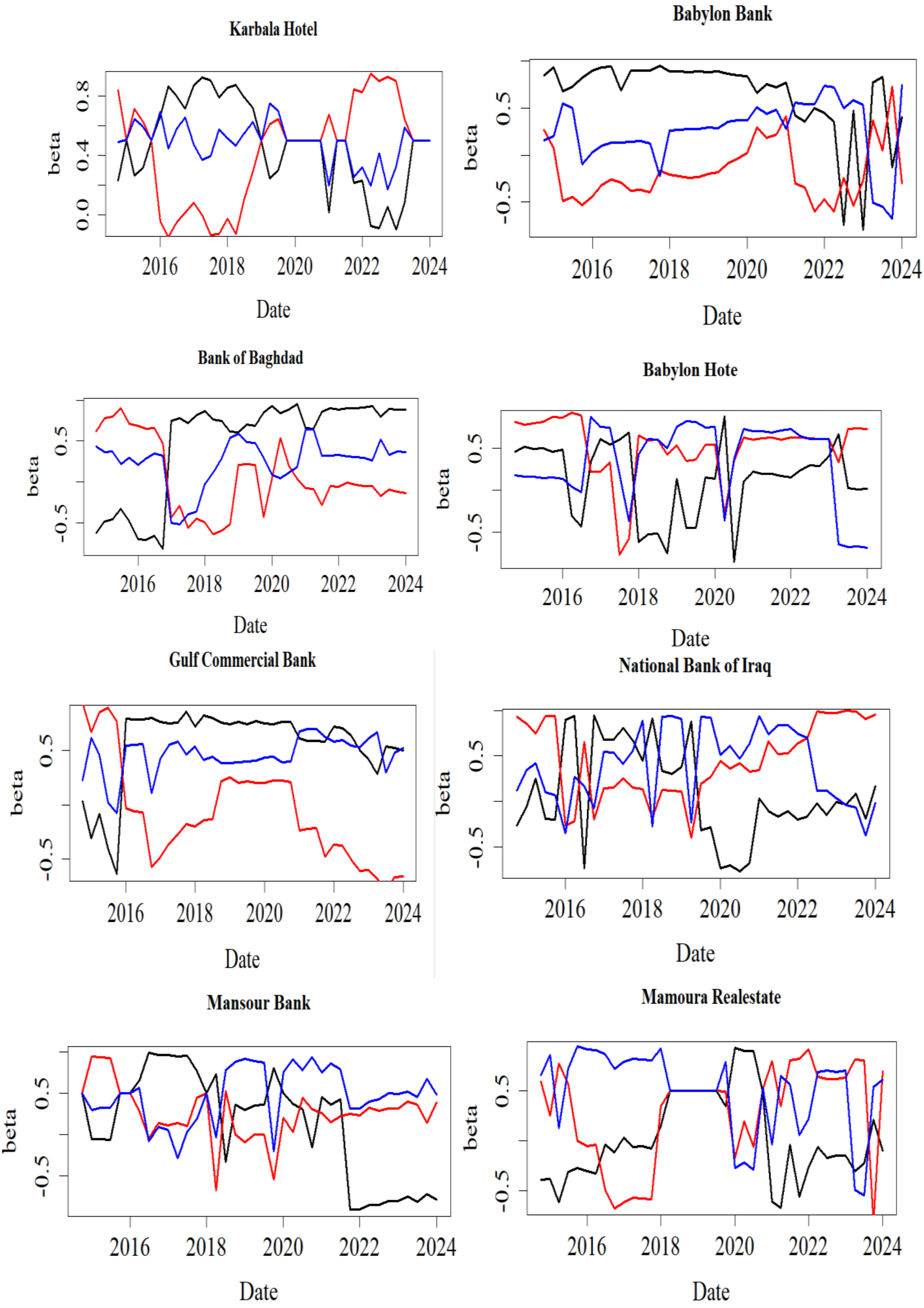
Our analysis examines the monitoring and assessment of changes in the beta coefficients over time for the three components of the Fama-French model, as illustrated in Figure (1). The beta of an asset captures its sensitivity to the hypothetical risks embodied by the three proxy variables, which are used as cross-sectional regression-independent variables. Instability of the coefficients is portrayed in time-varying beta plots illustrating how at different times the extent of the effects of the determinants on the volatility of assets varies. This form of analysis is important to comprehend how financial assets react to economic circumstances, which helps when making future investments.

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

Table 1: Model Fit Criteria

Name Of Asset	Statistic	KWR	SIMS	Statistic	KWR	SIMS
1. Babylon Bank	VR1 VR2	0.55708 0.59828	0.60461 0.16514	RMSE	0.42367	0.21215
2. Karbala Hotel	VR1 VR2	0.45827 1.19085	0.52380 0.74154	RMSE	0.10046	0.08593
3. Modern for Animal Production	VR1 VR2	0.28521 0.97279	0.20719 0.51355	RMSE	0.22738	0.18060
4. Babylon Hotel	VR1 VR2	2.28179 0.29925	0.96927 0.01350	RMSE	0.86049	0.18759
5. Mansour Hotel	VR1 VR2	0.23583 0.70904	0.39948 0.33965	RMSE	0.17073	0.12673
6. Commercial Bank of Iraq	VR1 VR2	1.07425 0.72453	0.73261 0.28557	RMSE	0.15065	0.09064
7. Bank of Baghdad	VR1 VR2	1.16575 0.73200	0.62850 0.29968	RMSE	0.17749	0.12391
8. Iraqi Middle East Bank	VR1 VR2	1.41136 0.79714	0.53764 0.39518	RMSE	0.22612	0.15136
9. investment Bank of Iraq	VR1 VR2	0.44158 0.41469	0.66147 0.21642	RMSE	0.12510	0.09562
10. National Bank of Iraq	VR1 VR2	0.61656 0.63402	0.32313 0.40051	RMSE	0.20925	0.17123
11. Sumer Commercial Bank	VR1 VR2	1.65454 0.96948	0.57386 0.37039	RMSE	0.08191	0.05766
12. Gulf Commercial Bank	VR1 VR2	0.82180 0.59345	0.53798 0.31472	RMSE	0.17190	0.11877
13. Mousl Bank	VR1 VR2	0.80033 1.05054	0.22537 0.54721	RMSE	0.30787	0.22169
14. Mansour Bank	VR1 VR2	0.42974 0.53882	0.25932 0.44928	RMSE	0.10649	0.09864
15. United Bank For Investment	VR1 VR2	1.14751 0.85486	0.72572 0.24386	RMSE	0.19740	0.11284
16. Al-Ameen for Insurance	VR1 VR2	1.08583 1.33078	0.42093 0.79387	RMSE	0.18271	0.17138
17. Mamoura Real-estate	VR1 VR2	0.47471 0.97733	0.46093 0.52208	RMSE	0.22460	0.16401
18. Al-Nukhba for Construction	VR1 VR2	0.73029 0.54700	0.36568 0.40894	RMSE	0.16075	0.12364
19. Modern Sewing	VR1 VR2	0.53895 0.94840	0.29316 0.59734	RMSE	0.20384	0.16956
20. Al-Mansour Pharmaceutical	VR1 VR2	0.48004 0.98920	0.21838 0.69334	RMSE	0.24483	0.21075
21. Baghdad Soft Drinks	VR1 VR2	0.60490 0.67218	0.38671 0.50677	RMSE	0.21431	0.17906
22. Al- Kindi of Veterinary	VR1 VR2	0.69580 0.81696	0.87369 0.10436	RMSE	0.24412	0.09568

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104



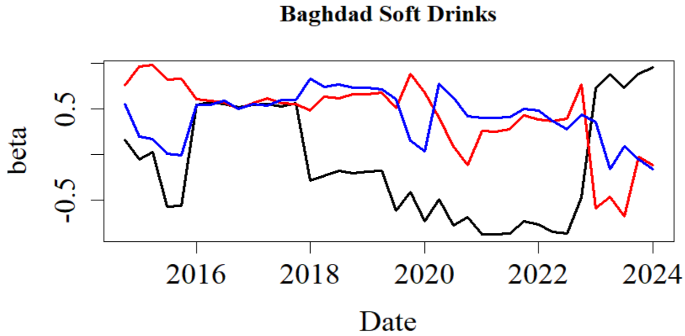


Figure 1: Beta Time Varying for Each Factor

The research emphasizes that Beta 1, which quantifies the degree to which a market factor influences excess returns, is different for different assets and may be higher or lower than one. A beta value larger than one signifies that an asset performs better than the market, with a rise in the return on the asset translating to a higher increase in market returns. Beta 2, which reflects the effect of company size on return, records both a positive and a negative correlation. Large companies, holding diversified assets, are likely to enjoy better market strengths and higher profitability, evidenced by a positive correlation. An inverse relationship, on the other hand, indicates that there are likely to be profitability issues for large companies and size-based weaknesses (Josefy et al., 2015). Beta 3 reflects the impact of the Value vs. Growth (HML) factor on returns, representing the historical excess returns of value stocks overgrowth stocks. This factor has shown both positive and negative correlations, with values consistently lower than one for each asset. The study concludes that asset behaviours differ over time, with economic and health conditions during the study period contributing to the instability of these coefficients.

The second step of the FM methodology involves cross-sectional regression to estimate the relationship between the response variable (excess returns on assets) and predictor variables (the time-varying betas of the Fama-French three-factor model) across a sample of assets at a given point in time. This process yields a series of gamma estimators, which help assess the performance of financial assets, analyse returns for the Fama-French factors, and test the significance of these influencing factors. The Fama-MacBeth model thus provides a robust analytical framework to understand the relationship between asset returns and market drivers. This aids in return analysis, key factor estimation, and stable investment decision-making.

It employs 28 cross-sectional estimation techniques based on KWR and SIMS to estimate gamma (risk premium) and test the statistical significance of the parameters, as presented on Table (2). The main purpose for using the Fama-MacBeth model is to provide an analytical tool with the potential to deal with the intricacies of asset return and market factors. By using the said model, the research performs statistical significance tests for each of the factors, estimating the size of its effect to see which

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

variables have the most impact on the dependent variable. This helps to establish whether the relationships seen between factors and returns are statistically significant or coincidental. By comprehending said estimations as well as their influencing factors, the investor can make sound, long-term investment choices.

The research made use of four different cross-sectional regression evaluation methods to fulfill its analytical goals. These methods relied on the KWR method and the SIMS, which were both applied for the purpose of time-varying betas estimation during the initial stage of research. The estimated gamma coefficients for each of the three Fama-French factors, as well as the findings from the respective significance tests for each method, are shown in Table (2), which summarizes the overall outcomes of the said estimates. By comparing the results from various approaches and assessing the sensitivity of the estimates to methodological changes, this robust approach enhances and corroborates the conclusions regarding the relationship between risk and return in the financial market.

The findings of a detailed analysis of the Fama-MacBeth model's parameters are presented in Table (2), which displays the average gamma (γ) coefficients for each asset's cross-section. Applying the same technique for estimating time-varying beta values (proxy variables) in the time series regression at the cross-sectional regression stage results in a notable convergence in the predicted gamma values.

Table 2: Test and Estimation of Parameters of Cross Sectional Regression

Estimate of Methods	Statistic	$\gamma_{0,t}$	$\gamma_{1,t}$	$\gamma_{2,t}$	$\gamma_{3,t}$
1. KWR-SIM	MEAN	0.248998	0.40289	0.25617	0.37568
	S.E	0.03034	0.05650	0.08209	0.06417
	T-STAT	8.20775	7.12965	3.12057	5.85402
	P Value	0.00000	0.00000	0.00343	0.0000
2. SIM-KWR	MEAN	-0.01584	-0.05560	0.02216	-0.01168
	S. E	0.02510	0.03013	0.03351	0.06312
	T-STAT	-0.63119	-1.84515	0.66124	-0.18502
	P Value	0.53169	0.07282	0.51245	0.85420
3. SIM –SIM	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
4. KWR-KWR	MEAN	-0.02776	-0.01202	0.01038	-0.03895
	S.E	0.01591	0.02469	0.03778	0.03530
	T-STAT	-1.74377	-0.48682	0.27489	-1.10333
	P Value	0.08928	0.62917	0.78488	0.27682

This consistency suggests that, despite slight variations in the estimation methods, the core risk-return relationships remain stable. However, assessing the statistical significance of the computed gamma coefficients is essential to determine the validity

Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) (**pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

and relevance of these findings. Financial economic theories generally assume that the variables influencing asset return variations must be statistically significant. It is important to recognise that while statistical significance is critical in evaluation, it is not an absolute standard, particularly when comparing results across markets or countries with different economic characteristics. Distinct institutional frameworks, market growth stages, and economic structures can influence the relationship between risk and return, thus impacting the statistical significance of the estimated parameters.

As seen from the results in Table (2), the statistical relevance of the estimated parameters differs considerably based on the method applied for estimation. The SIMS-based technique showed statistically significant factors for asset returns at the initial estimate of beta as well as at the gamma estimation step. This indicates that from the given data and the prevailing socioeconomic situation, the SIMS model better reflects the expected risk-return relationship. As opposed to that, KWR technique, when implemented at both steps, did not provide statistically relevant parameters. This is because of the statistical properties of the KWR technique, where the findings were identical when using the same technique for both beta estimation for the time series regression and gamma estimation for the cross-sectional regression (Ngu et al., 2024).

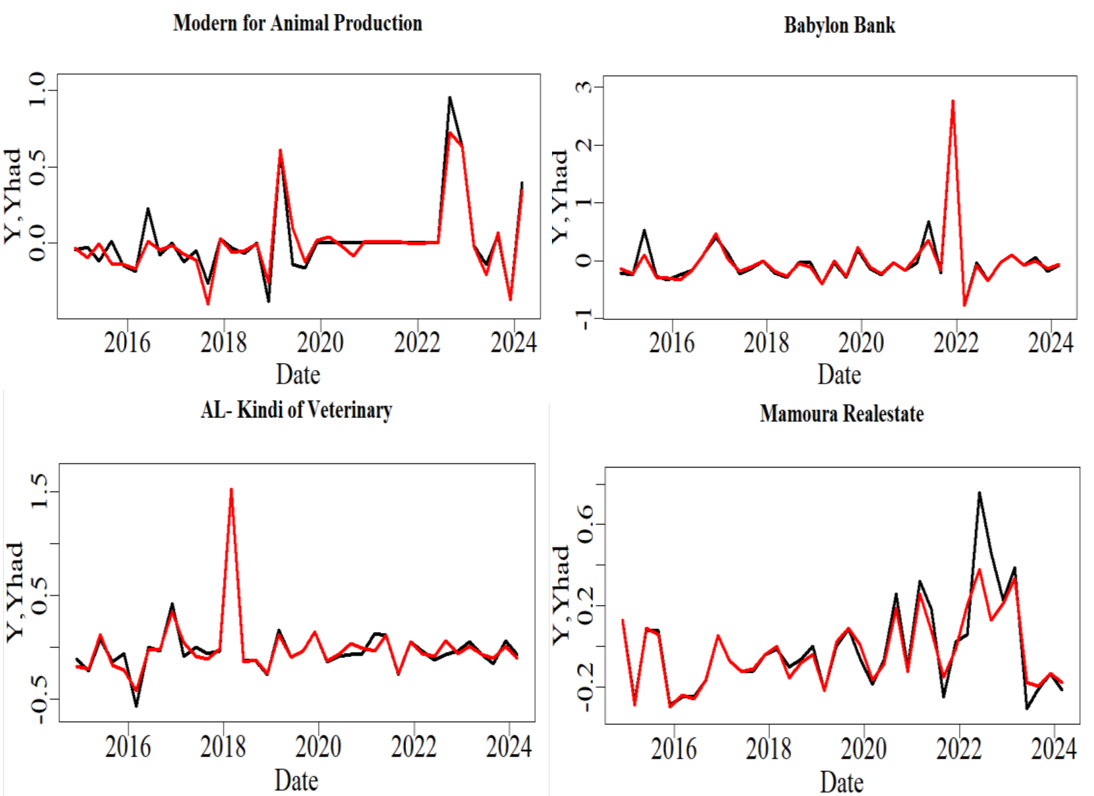
While statistical significance is a determining fact when assessing results, it is relative, particularly with the different economic contexts of states. The approach of SIMS indicated the importance of the determining factors, while KWR indicated parameter insignificance. The research used the RMSE indicator to find the optimal statistical estimator, because RMSE has the capability to give a clear, unambiguous representation of average estimation errors as is, using the original units of measurement. Since RMSE is a representation of average absolute error of predictions with respective actual values, it helps understand results. Also, with its penalization of extreme errors, RMSE is extremely responsive to the overall performance of the estimator, as well as its capability to prevent implausible estimates. The outcomes of the metric for each estimator are indicated in Table (3).

Table 3: ARMSE of Cross Sectional

Method	ARMSE
1. KWR-SIMS	0.08152
2. SIMS-KWR	0.24676
3. SIMS-SIMS	0.12842
4. KWR-KWR	0.26109

Table (3) illustrates the results of each estimation method of the RMSE values, with the KWR-SIMS method being a clear outperformer compared to the other methods. The KWR-SIMS method recorded the smallest RMSE, thus it is the most accurate method for creating forecasts with the least error. This reflects its efficiency in narrowing the estimated and actual difference, and thus it is the best method for use with the current study. The KWR-SIMS technique is depicted even further by Figure (2) where actual vs. forecasted return relations from cross-sectional regression analysis are illustrated. The effectiveness of the KWR-SIMS technique is seen as the calculated expected excess return tracks closely with actual excess return.

This graphical depiction makes it easy to compare the theoretical predictions with actual figures, illustrating the capability of the method to minimize pricing errors. By analyzing the parallelism of the lines for actual and anticipated returns, the research can better gauge how well the Fama-MacBeth model, used with the KWR-SIMS method, is able to reduce differences in asset return predictions. Figure (2) is a crucial graphical representation, with the excess return paths being very much aligned with the excess return from our model for a cross-section of listed companies. This alignment across the times tested in cross-sectional regression is evidence of the robustness and reliability of the representation of the model. Of keen observation is that the model is also able to detect major oscillations and random jumps from actual return series, which is proof of the strength of the model approximating actual dynamics.



Vol: 17 **No:** 01 **Year:** 2025 **ISSN:** 1309-8055 (Online) **(pp.** 46-64) **Doi:** 10.34109/ijefs.202517104

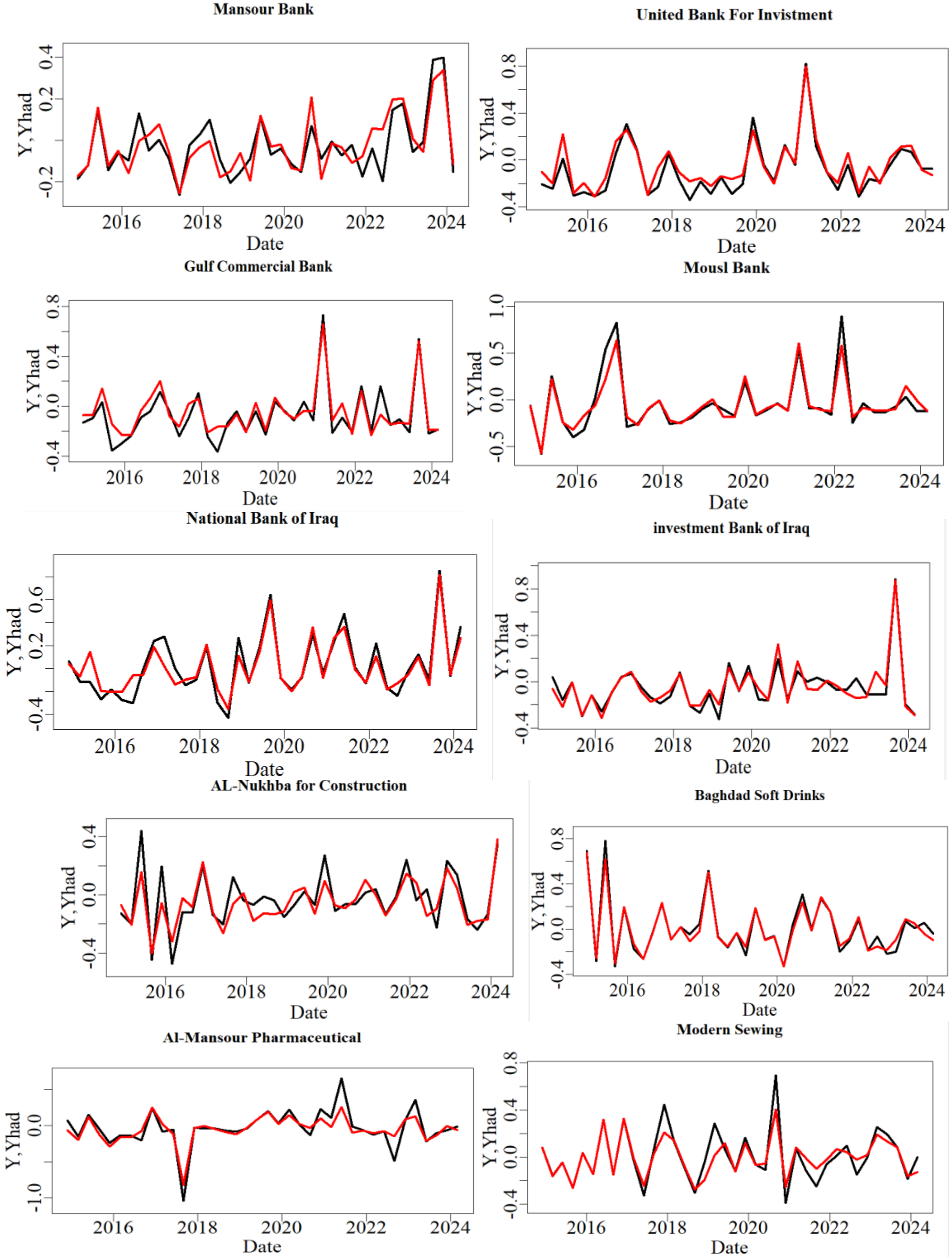


Figure 2: Plot of Realized Excess Return versus Predicted Excess Return in Cross Section Regression by KWR -SIM

CONCLUSION

This study applied Fama-MacBeth regression and the Fama-French three-factor model to analyse asset returns in the Iraq Stock Exchange, focusing on 22 companies across six sectors from Q1 2010 to Q1 2024. We evaluated dynamic beta coefficients using the SIM and KWR, with SIM consistently outperforming KWR in model fit and prediction precision, achieving a lower average RMSE of 0.14316. Time-varying beta visualisations revealed significant fluctuations in Beta 1, Beta 2, and Beta 3, reflecting economic and health changes during the study period. The SIM-SIM approach in cross-sectional regression showed statistical significance across all factors, with t-statistics of 3.37, 3.93, and 9.37 for Beta 1, Beta 2, and Beta 3, respectively. This improved asset and portfolio assessment, providing better risk management and pricing accuracy, offers valuable insights for financial analysts and investors.

REFERENCES

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. https://doi.org/10.22075/ijnaa.2022.25984.3180

Ang, A., & Kristensen, D. (2012). Testing conditional factor models. *Journal of Financial Economics*, *106*(1), 132-156. https://doi.org/10.1016/j.jfineco.2012.04.008

Aydına, D., Günerib, Ö. İ., & Fitc, A. (2016). Choice of Bandwidth for Nonparametric Regression Models using Kernel Smoothing: A Simulation Study. *International Journal of Sciences: Basic and Applied Research*, *26*(1), 47-61. https://core.ac.uk/download/pdf/249335133.pdf

Bryzgalova, S., Huang, J., & Julliard, C. (2024). Bayesian Fama-MacBeth Regressions. *Available at SSRN*. https://doi.org/10.2139/ssrn.4989615

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The journal of finance*, *50*(1), 131-155. https://doi.org/10.1111/j.1540-6261.1995.tb05169.x

Ferreira, E., Gil-Bazo, J., & Orbe, S. (2011). Conditional beta pricing models: A nonparametric approach. *Journal of Banking & Finance*, *35*(12), 3362-3382. https://doi.org/10.1016/j.jbankfin.2011.05.016

Ferson, W. E., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of political economy*, *99*(2), 385-415. https://doi.org/10.1086/261755

Fraser, P., Foort, H., Martin, H., & and Macgregor, B. (2004). Time-varying betas and the cross-sectional return–risk relation: evidence from the UK. *The European Journal of Finance*, *10*(4), 255-276. https://doi.org/10.1080/13518470110053407

González, M. V. E., & Mandaluniz, S. O. (2006). Nonparametric estimation betas in the

Market Model. *Documentos de Trabajo BILTOKI*(3), 1. https://dialnet.unirioja.es/servlet/articulo?codigo=2138580

Green, J., Hand, J. R. M., & Zhang, X. F. (2016). Errors and questionable judgments in analysts' DCF models. *Review of Accounting Studies*, *21*(2), 596-632. https://doi.org/10.1007/s11142-016-9352-4

Han, Y., He, A., Rapach, D. E., & Zhou, G. (2024). Cross-sectional expected returns: new Fama–MacBeth regressions in the era of machine learning. *Review of Finance*, *28*(6), 1807-1831. https://doi.org/10.1093/rof/rfae027

Härdle, W. (2004). *Nonparametric and semiparametric models*. https://doi.org/10.1007/978-3-642-17146-8

Haslam, S. A., Steffens, N. K., Cruwys, T., & McGarty, C. (2024). Research methods and statistics in psychology. https://www.torrossa.com/it/resources/an/5909100

Henley, S. S., M., G. R., & and Kashner, T. M. (2020). Statistical modeling methods: challenges and strategies. *Biostatistics & Epidemiology*, *4*(1), 105-139. https://doi.org/10.1080/24709360.2019.1618653

Hodoshima, J., Garza–Gómez, X., & Kunimura, M. (2000). Cross-sectional regression analysis of return and beta in Japan. *Journal of Economics and Business*, *52*(6), 515-533. https://doi.org/10.1016/S0148-6195(00)00031-X

Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development Discussions*, *2022*, 1-10. https://doi.org/10.5194/gmd-15-5481-2022

Hur, J. (2024). A Cross-sectional Asset Pricing Test with More Power: An Instrumental Variable Approach. In *Handbook of Investment Analysis, Portfolio Management, and Financial Derivatives: In 4 Volumes* (pp. 3547-3581). World Scientific. https://doi.org/10.1142/9789811269943_0108

Jagannathan, R., & Wang, Z. (1998). An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *The journal of finance*, *53*(4), 1285-1309. https://doi.org/10.1111/0022-1082.00053

Josefy, M., Kuban, S., Ireland, R. D., & Hitt, M. A. (2015). All things great and small: Organizational size, boundaries of the firm, and a changing environment. *Academy of Management Annals*, *9*(1), 715-802. https://doi.org/10.5465/19416520.2015.1027086

Karolyi, G. A., & Van Nieuwerburgh, S. (2020). New methods for the cross-section of returns. *The Review of Financial Studies*, *33*(5), 1879-1890. https://doi.org/10.1093/rfs/hhaa019

Köhler, M., Schindler, A., & Sperlich, S. (2014). A review and comparison of bandwidth selection methods for kernel regression. *International Statistical Review*, *82*(2), 243-274. https://doi.org/10.1111/insr.12039

Liu, W., & Huang, J. Z. (2021). A New Model of Capital Asset Prices. https://doi.org/10.1007/978-3-030-65197-8

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

Ngu, J. C. Y., Yeo, W. S., Thien, T. F., & Nandong, J. (2024). A comprehensive overview of the applications of kernel functions and data-driven models in regression and classification tasks in the context of software sensors. *Applied Soft Computing*, *164*, 111975. https://doi.org/10.1016/j.asoc.2024.111975

Patra, R. K. (2016). *Semiparametric inference with shape constraints*. Columbia University. https://www.proquest.com/openview/fb1dd7491ac3d56511f27a9407b57305/1?pq-origsite=gscholar&cbl=18750

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics*, *60*(1), 13-50. https://doi.org/10.1007/s00181-020-01875-7

Qadri, S. S. S. M., Almusawi, A., Albdairi, M., & Esirgün, E. (2024). Optimizing Traffic Signal Timing at Urban Intersections: A Simheuristic Approach Using GA and SUMO. 2024 Innovations in Intelligent Systems and Applications Conference (ASYU), https://doi.org/10.1109/ASYU62119.2024.10757086

Škrinjarić, T. (2018). Rolling regression CAPM on Zagreb stock exchange: Can investors profit from it? *Economic Review: Journal of Economics and Business*, *16*(2), 7-22. https://www.econstor.eu/handle/10419/267868

Taheri, S., Hosseini, P., & Razban, A. (2022). Model predictive control of heating, ventilation, and air conditioning (HVAC) systems: A state-of-the-art review. *Journal of Building Engineering*, *60*, 105067. https://doi.org/10.1016/j.jobe.2022.105067

Toimil, A., Losada, I. J., Nicholls, R. J., Dalrymple, R. A., & Stive, M. J. F. (2020). Addressing the challenges of climate change risks and adaptation in coastal areas: A review. *Coastal Engineering*, *156*, 103611. https://doi.org/10.1016/j.coastaleng.2019.103611

Wu, W., Chen, J., Yang, Z., & Tindall, M. L. (2021). A cross-sectional machine learning approach for hedge fund return prediction and selection. *Management Science*, *67*(7), 4577-4601. https://doi.org/10.1287/mnsc.2020.3696

Yu, Y., & and 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