Luis Ferruz Agudo y Guillermo Badía Fraile

Model adapted from Fama and French, paying attention to the assumptions of the estimation methodology

Modelo adaptado de Fama y French, prestando atención a las asunciones de la metodología de estimación

ABSTRACT

We examine whether the Fama and French (1992) model can be adapted to become a more versatile and flexible tool, capable of incorporating variations of company characteristics in a more dynamic form. For this, the risk factors are reconstructed at the end of each reading of monthly data. Over time, the evaluation of a company may change as a result of variations in, for example, its size, and we are aware that the Fama & French model does not accurately reflect these dynamics. Our purpose in this paper is also to highlight the need to analyze the estimation methodology assumptions of the model coefficients so as to avoid potential biases and inconsistencies. Our results show that the model appears to be statistically significant for the majority of the stocks and that the estimation methodology hypotheses are maintained, with some exceptions.

Keywords: Financial Models, Fama and French Model, Ordinary Least Square (OLS), Management Strategy.
JEL Classification: G11

RESUMEN

En el artículo se estudia si es posible adaptar el modelo de Fama y French (1992) para capturar de forma más dinámica las variaciones en las características de las compañías. Para ello se reconstruyen los factores de riesgo al final de cada mes tomando datos mensualmente. En el trascurso de un año, pueden cambiar las características de las compañías, por ejemplo, su tamaño, lo cual, mediante el procedimiento que siguen Fama y French

Department of Accounting and Finance, University of Zaragoza, Spain. lferruz@unizar.es gbadiaf@unizar.es

Departamento de Contabilidad y Finanzas, Universidad de Zaragoza, España.

Acknowledgements
The authors wish to state that the financing of this research was supported by the Government of Aragón and by the EUROPEAN SOCIAL FUND, budgets: G/17030/5423/440030/11101 (50%) y 17030/5423/440030/91101 (50%). We also wish to acknowledge the technical and operational support provided by GIECOFIN. It goes without saying that any errors are the sole responsibility of the authors.

Luis Ferruz y Guillermo Badía:
Model adapted from Fama and French, paying attention to the assumptions of the estimation methodology
Modelo adaptado de Fama y French, prestando atención a las asunciones de la metodología de estimación
***Análisis Financiero* n° 130. 2016. Págs.: 24-39**

(1992), no es reflejado hasta finales de año. Asimismo, se pretende resaltar la necesidad de analizar las asunciones del método de estimación de los coeficientes del modelo para que las inferencias que se realizan sean correctas. Los resultados indican que el modelo aparece estadísticamente significativo para la inmensa mayoría de títulos y que las hipótesis del método de estimación se mantienen salvo para determinados títulos.

Palabras clave: Modelos financieros, Modelo de Fama y French, Mínimos cuadrados ordinarios, Estrategias de inversión.
Códigos JEL: G11

Recibido: 18 de enero de 2016 **Aceptado: 2 de febrero de 2016**

1. INTRODUCTION

Financial assets pricing models are an essential tool for money managers. Since Sharpe (1964), Lintner (1965) and Black (1971) developed the Capital Asset Pricing Model (CAPM), a range of methodologies have appeared to determine the expected value of an asset. A model of the statistical type that evolved out of the CAPM was proposed by Fama and French (1992). Not only does this model incorporate the market as a risk factor, as the CAPM does, but it includes the size of a company and whether the company has value or is experiencing growth, as measured by the book-to-market ratio.

This new model, also known as the three-factors model, has been analyzed and evaluated often in the literature and is still a topic of debate, since there is no single conclusion concerning its reliability as a tool for estimating the yields of securities, and as an instrument of portfolio management.

Authors such as Miles and Timmermann (1996), Fletcher (2001), Hussain et al. (2002), Hwang and Satchell (2005), Lajili-Jarjir (2007), Agarwal and Poshakwale (2010), Hamard and Mascareñas (2010), Abhakorn et al. (2013), Eraslan (2013), Gregory et al. (2013), Nichol and Dowling (2014) Zhong et al. (2014) and Ferruz and Badía (2015), among many others, have analyzed the ability of the Fama & French model (F&F, henceforth) to explain the behavior of portfolios and individual securities in different markets, and under various assumptions.

In this paper, we pay attention to the procedure that Fama and French (1992) follow to build risk factors and to construct the value and size of portfolios. The risk factors are constructed from data taken once a year and the value and size of portfolios are assessed once a year, maintaining invariability during the whole period. However, it should not be forgotten that variations can occur in the characteristics of a company during a given 12-month period, which will not be accounted for by the F&F procedure.

We test the capacity of the model by taking month-to-month data and rebuilding the value and size portfolios at the end of each month with the aim of developing a model that is more dynamic and adaptable.

This approach has two clear implications. First, the BM ratio varies according to the characteristics of the company at any given moment. Although we can expect that the numerator or book price appears invariant for an entire period, the denominator or market price does vary and, therefore, a company may be moving between different value portfolios during the year without being captured, i.e. between Low, Medium and High portfolios. Second, our approach provides greater variability in the size factor, which allows us to capture variations in the capitalization of the company as a result of, among other things, price fluctuations of the stock, which can have consequences for the classification of a company (such as, Big or Small).

The variations that may arise in the characteristics of a company, causing it to move between different value and size portfolios, has a direct impact on the associated return of the portfolios. Therefore, our approach allows us to form new portfolios every month by utilising the true set of characteristics, and the yields reflected by these portfolios are better suited to the situation at a given moment.

To test the ability of the adapted model, the significance of the model as a whole, and the individual coefficients considered in the regression are checked.

Our second purpose is to demonstrate the need for, and importance of, an examination of certain aspects of the methodology to estimate the model coefficients.

Many studies have examined the capacity of the F&F model, including Bartholdy and Peare (2005), Pettengill et al. (2012) and Fama y French themselves (1992, 1993, 1996), among others, and conclusions have been reached without paying attention to the procedure underlying the estimation of the coefficients, i.e. without testing the degree of compliance with the assumptions of the estimation methodology, in this case, Ordinary Least Square (OLS).

It must be remembered that, if the estimated parameter is biased or inconsistent, incorrect estimates will result. Under these circumstances, a manager who is implementing an investment strategy based on inferences made by the model, may very well find that the decisions he takes will lead to unexpected results.

Thus, after determining the capacity - or incapacity - of the adapted F&F model, we analyze the degree of compliance with the hypothesis of the estimation methodology of OLS.

Our purpose, then, is to present an adequate and efficient model for making investment decisions, which will allow better adaptation to the environment. In any case, whether or not our results are satisfactory, we highlight the need to analyze the level of compliance with the hypothesis of the estimation methodology, since if this is not performed, the inferences drawn will be inefficient and inaccurate.

The rest of the paper is organized as follows. In Part Two we present the data and period, the model used, and the assumptions of the estimation methodology. In Part Three, we offer the results of the empirical analysis, and Part Four contains a discussion of our conclusions.

2. DATA AND METHODOLOGY

Our data covers the period from January, 2006 to December, 2010. Authors such as Chen et al. (2014) also take a five-year interval to estimate the coefficients. This period, as argued by Brooks (2014), is often used for this purpose, and Bartholdy and Peare (2005) carry out a study that determines that this is the optimal period for to estimate the coefficients.

We sample a total of 692 non-financial firms trading in the UK[1] Market. Company prices are taken from the Morningstar data base. To qualify for our study, these securities must have been in the UK market for the whole period under analysis (this is why we assume a certain survival bias).

The FTSE All Share Index is used as a proxy for the market portfolio, so as to accurately cover the joint realistic opportunities of investment that appear to investors.

Following Nichol and Dowling (2014), we have selected the 3-month UK Treasury Bill rate as the risk-free asset.

For returns, we have used monthly prices, thus limiting the problem that certain stocks do not necessarily trade on every single day. The above-mentioned profitabilities are calculated as the natural logarithm of the quo-

[1] Financial companies are excluded, for the same reasons as Fama and French (1992).

tient between the closing price for the analyzed month and the closing price for the previous month.

2.1. Building the Adapted Model

In view of our first aim, we reconstruct the risk factors at the end of each month, from the monthly data. Thus, to get the size factor we take the stock exchange capitalization at the end of each month, whereas the book-to-market (BM) factor is obtained as the quotient between the book price and the share market-price, both taken at the end of every year.

In contrast, in the original model, Fama and French (1992) take the value of market capitalization in June to calculate the size factor, and the book price at the end of the year, along with the share price on the last day of December of the given year, to obtain the BM ratio.

Once the risk factors have been obtained, we proceed to construct the SMB (Small minus Big) and HML (High minus Low) portfolios. The procedure is the same as that followed by Fama and French (1993), except that they build their portfolios once a year (in June) and hold them during the entire period, whereas we conduct our procedure on a month-to-month basis.

Thus, to construct the SMB portfolio, we rank the securities by capitalization value at the end of each month and establish two groups, dividing the sample by the median. In this way, we have the large-capitalization assets on one side, and the small-capitalization assets on the other.

Subsequently, every group is ranked from highest to lowest according to the B/M value and divided into three subgroups, taking the same values as Fama and French (1992), with the percentiles of 30 and 70 percent. In this way, the Value portfolio (High), the Neutral portfolio (Medium) and the Growth portfolio (Low) of Big and Small companies are obtained.

The SMB risk factor is the average return associated with the difference between the average return portfolio of small-cap companies (S/L, S/M, S/H) and the mean return portfolio of large-cap companies (B/L, B/M, B/H), whereas the HML factor risk is the average return portfolio of Value assets (S/H, B/H), minus the average return portfolio of Growth assets (S/L, B/L).

In this way, we create six portfolios, by size and BM ratio, and reconstructed each month according to changes in the characteristics of the companies, with the target being to provide greater reaction and adaptability to the model under certain contingencies.

Fama and French (1992) considered 100 portfolios to test their model, while authors such as Rogers and Securato (2007), Grauer and Janmaat (2010), Abhakorn et al. (2013) and Fama and French (1993) use 25. Since our first objective is to analyze the capacity of the model to fit the returns of individual securities, the model is run with this purpose on each asset considered. Avramov and Chardia (2006) and Soumaré et al. (2013) point out that using individual stock rather than portfolios avoids the loss of information when assets are grouped into portfolios, and limits biases in the data associated with portfolio construction.

Once the portfolios SMB and HML are obtained for each month, time series regressions are run for each security in which the coefficients of the model are estimated.

$$R_i - R_f = \alpha_i + \beta_i RMRF + \beta_i SMB + \beta_i HML + \varepsilon_i \quad (1)$$

Where *Ri* is the performance of the securities, *Rf* is the return on the risk-free asset, *RMRF* is the Market risk factor, i.e. the excess return of the Benchmark on the risk free asset, *SMB* is the difference between the Small-cap stocks portfolio returns and the Large-cap stocks portfolio returns, and *HML* is the difference between the High securities portfolio returns and the Low securities portfolio returns.

Provided that the model fits the returns of the stocks, it will achieve a minimization of the sum of the squared areas painted from each point obtained from the relation profitability-risk factors to the line drawn by the model, and it will adjust the estimated returns of the model to the truly-obtained ones for the stocks.

The procedure to estimate the coefficients is Ordinary Least Square (OLS)[2]. As mentioned above, in order for the inferences achieved from the model to be accurate, a set of prerequisites must be maintained.

2.2. Statistics to test the ability of the Adapted Model

Two different tests are employed to check the capacity of the model: *t-statistic* and *F-statistic*.

The *t-statistic* measures the statistical significance of each of the coefficients in the set of the regression. The Null Hypothesis is that the coefficients are equal to zero, i.e. $\alpha_i = 0$, $\beta RMRF = 0$, $\beta SMB = 0$, $\beta HML = 0$. In this particular case the *t-statistic* is achieved as follows:

$$t - statistic = \frac{\hat{\beta}}{SE(\hat{\beta})} \quad (2)$$

Where $\hat{\beta}$ is each of the estimated coefficients and $SE(\hat{\beta})$ are the standard errors of such coefficients. The expression to obtain the intercept standard errors is the following:

$$SE(\hat{\alpha}) = \sigma^2 \sqrt{\frac{\sum x_t^2}{T((\sum x_t^2) - T\bar{x}^2)}} \quad (3)$$

In order to calculate the beta standard error:

$$SE(\hat{\beta}) = \sigma^2 \sqrt{\frac{1}{\sum x_t^2 - T\bar{x}^2}} \quad (4)$$

The standard errors give a measure of the uncertainty level of the estimated values for the coefficients. These are a function of the explanatory variable *x*, the sample size *T* and the variance σ^2.

The *F-statistic* measures the capacity of the coefficients of simultaneous way, that is, the model as a whole. This test works with two regressions, a restricted regression and another unrestricted. The coefficients of the unrestricted regression are freely determined with the data; in other words, it is the regression achieved in the previous step. The restricted regression is the one in which the coefficients are restricted, that is, the restrictions are imposed on the coefficients.

In our case, we wish to test whether the model is statistically different to zero as a whole, therefore, the restricted regression is:

$$R_i - R_f = \alpha_i + \beta_i RMRF + \beta_i SMB + \beta_i HML + \varepsilon_i \quad (5)$$

Subject to: $\alpha_i + \beta_i RMRF + \beta_i SMB + \beta_i HML = 0$

To run the statistic, it is necessary to determine the residual sum of the square of each of the regressions and then compare them. The expression of the test is as follows:

$$F - statistic = \frac{RRSS - URSS}{URSS} \, x \, \frac{T-k}{m} \quad (6)$$

Where *URSS* is the Residual Sum of Squares from Unrestricted regression $\Sigma_t(y_t - \hat{y}_t)^2$, *RRSS* is the Residual Sum of Squares from Restricted regression, *m* is the restrictions number, *T* is the observations number, and *k* is the regressors number in the unrestricted regression, including the constant.

2.3. Assumptions by estimation method

In several studies, the efficiency of the F&F model is analyzed via certain statistical measures, such as the ones used here, with the *rolling window* proposed by Fama and MacBeth (1973), or with performance measures such as Jensen's *alpha* (1968) or the Treynor ratio (1973); and with these results, conclusions are drawn.

However, as mentioned earlier, for the estimations carried out by the F&F model to be reliable, and so that the

[2] The mathematical framework behind obtaining each of the beta coefficients and the intercept can be found in Gujarati (2003) and Brooks (2014), among others.

innovation can safely be used by managers, certain aspects of the methodology do have to be re-examined.

From a Classical Linear Regression Model (CLRM) estimated through OLS of the type

$$y_t = \alpha + \beta x_t + \mu_t \quad t = 1, 2, \ldots, T \qquad (7)$$

Sharpe (1964), Lintner (1965), and Black (1972) developed the Capital Asset Pricing Model (CAPM), based on the two fundamental parameters, *risk* and *return*, intended to first achieve a complete equilibrium of the capital market.

Nevertheless, even though the development of this model represented an undoubted advance, this itself has been widely criticized, due to the fact that the systematic risk factor is not sufficient to explain the expected return.

Subsequently, Fama and French (1992), taking as starting point the CAPM, and after checking those aspects that significantly affect the behavior of stocks in capital markets, built a model of a statistical type that considered a new set of factors.

The authors based their new approach on the same methodology to estimate the parameters, i.e. OLS, but, because of the incorporation of new elements in the regression, the earlier model is inadequate and generates a regression with several coefficients to be estimated, i.e. it generalizes the simple model in a multiple linear regression model, with the following shape:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_k x_{kt} + \mu_t \quad t = 1, 2, \ldots, T \qquad (8)$$

For the estimated coefficients to be consistent, unbiased, and with the inferences performed correctly, various properties must be maintained: the errors have zero mean, the variance of the errors is constant and finite over all values of x_t, the errors are linearly independent of one another, there is no relationship between the error and the corresponding x variable, and the error term is normally distributed.

As long as the first four assumptions are maintained, the estimated parameters by OLS meet the desirable properties, otherwise known as the Best Linear Unbiased Estimators (BLUE). The last property, as will be seen later, is not necessarily exclusive, owing to the Central Limit Theorem.

Different tests are used to check the compliance of each of these assumptions.

The errors have zero mean: $E(\mu_t) = 0$

The first requirement is always satisfied if the calculated regression includes a constant term. When the regression does not include a constant, and the average value of the errors is not zero, the coefficient values could be biased because the regression line would be forced through the origin.

The variance of the errors is constant and finite over all values of x_t: var $(\mu_t) = \sigma^2 < \infty$

This is known as the homoskedasticity assumption. If the errors do not have constant variation, they are said to be heteroskedastic.

There are several tests to determine that the residuals of the model are not heteroskedastic. Goldfeld and Quandt (1965) developed a process which divided the total sample into two sub-samples, estimating the model on each sub-sample, calculating the residual variances and verifying that they are equal for both sides. The construction of this test is simple, but the conclusions may not be entirely acceptable, because the point at which the sample is split, is taken arbitrarily and this can cause the results to differ if another point is taken.

A more general test is proposed by White (1980), where the model is first estimated with the following general form:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t \qquad (9)$$

Then, the residuals are obtained and a squared regression is run, containing the original explanatory variables, the square of these variables, their cross-products, and a constant.

$$\hat{u}_t^2 = \alpha_1 + \alpha_2 x_{2t} + \alpha_3 x_{3t} + \alpha_4 x_{2t}^2 + \alpha_5 x_{3t}^2 + \alpha_6 x_{2t} x_{3t} + v_t \quad (10)$$

Where v_t is a normally distributed disturbance term independent of υ_t.

To obtain the auxiliary regression, the test can be performed using the F-test previously discussed. Thus, on the one hand, an unrestricted auxiliary regression is run, and on the other, an auxiliary regression is run with the restrictions previously imposed, i.e. the sum of all coefficients is equal to zero. When both regressions have been obtained, the RSS of each one is used as input in the F-test formula.

The errors are linearly independent of one another $cov(\mu_t, \mu_j) = 0$

The estimated model must have zero covariance between the error terms over time, i.e. the autocorrelation in the model errors is analyzed.

There are several tests to check the existence of autocorrelation. Durbin-Watson (1951) develop a test for first order autocorrelation, testing the autocorrelation between an error and its immediate previous value. Given the limitation of this test, since relationships can exist that are far apart, the Breusch-Godfrey test is used, which is a more general test for autocorrelation by which it is possible to check higher order relationships. The model for errors presents the following form:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \cdots + \rho_r u_{t-r} + v_t \qquad v_t \sim N(0, \sigma_v^2) \quad (11)$$

Where H_0: $\rho_1 = 0, \rho_2 = 0 \ldots \rho_r = 0$

and H_1: $\rho_1 \neq 0, \rho_2 \neq 0 \ldots \rho_r \neq 0$

So, under the null hypothesis, the current error is not related with any of its r previous values. The test is performed in the following way:

To begin, the model is estimated, and the residuals are obtained, $\hat{\upsilon}_t$. Subsequently, we run a regression of the residuals, with all the regressors considered, as before, as well as the necessary lags.

The general expression is of this type:

$$\hat{u}_t = \gamma_1 + \gamma_2 x_{2t} + \gamma_3 x_{3t} + \gamma_4 x_{4t} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \rho_3 \hat{u}_{t-3} + \cdots + \rho_r \hat{u}_{t-r} + v_t \qquad v_t \sim N(0, \sigma_v^2) \quad (12)$$

A difficulty with this test is determining the number of lags, r, of the residuals that are incorporated in the auxiliary regression. One possibility, used in prior literature, is take the data frequency, i.e. since the monthly observations are employed, it is determined that r is equal to 12. Thus, it is expected that the errors do not have a higher relationship than the annual period.

Subsequently, we obtain the R^2 of this auxiliary regression of the errors and, let T as the number of observations, the statistic test is given by:

$$(T - r)R^2 \sim X_r^2 \quad (13)$$

There is no relationship between the error and the corresponding x variable: $cov(\mu_t, \mu_t) = 0$

The following hypothesis points out that the explanatory variables are deterministic or not of a stochastic nature. Violation of this restriction is known as the presence of stochastic regressors.

Nevertheless, the OLS estimator is consistent and unbiased in the presence of stochastic regressors, as long as they are not correlated with the error term. This can be seen as follows:

Given the mathematical expression to obtain the beta coefficients of the regression, and the expression to obtain the explained variable

$$\hat{\beta} = (X'X)^{-1}X'y \quad (14)$$

and

$$y = X\beta + u \quad (15)$$

Operating

$$\hat{\beta} = (X'X)^{-1}X'(X\beta + u) \quad (16)$$

$$\hat{\beta} = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u \quad (17)$$

$$\hat{\beta} = \beta + (X'X)^{-1}X'u \quad (18)$$

If expectations are taken, providing that X and u are independent:

$$E(\hat{\beta}) = E(\beta) + E((X'X)^{-1}X'u) \quad (19)$$

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X']E(u) \quad (20)$$

Due to the fact that $E(u) = 0$, this expression takes zero value and therefore, even if the regressors are stochastic, the estimation is unbiased. On the other hand, if it is found that one or more explanatory variables is contemporaneously correlated with the disturbance term, the OLS estimator will not be consistent. So, the variables would have greater explanatory power, when in fact this would be a consequence of the correlation between the error and the explanatory variable. As a result, biased and inconsistent estimated parameters would result and inefficient estimates would be generated.

Therefore, the interest is to analyze the relationship between the variables and the disturbance term.

The disturbances are normally distributed: $\mu_t \sim N(0,\sigma^2)$
The normality assumption is the final property, to determine that the errors of the model are normally distributed. Since the explained variable depends partially on the disturbance term, it is possible to state that if the error is normally distributed, this variable will be too.

If normality is not found, a method can be used that does not assume it, but this is complicated to implement and, in any case, the properties of the method could be less stable. However, for sample sizes that are sufficiently large, the fulfillment of the normality assumption is not determinant. According to the Central Limit Theorem, the statistics tests follow a suitable distribution, even in the absence of error normality. This theorem states that the sample distribution of any set of random observations tends towards the normal distribution with mean equal to the population mean, as the sample size tends to infinity.

3. EMPIRICAL ANALYSIS

3.1. Adapted Model ability evaluation

First, the results obtained by the *t-statistic* are presented. Since there are a large number of stocks, Table 1 shows the number of times that each of the coefficients appears as statistically significant in each of the regressions run for each stock. For instance, if the *p-value* associated with the *t-statistic* of the RMRF coefficient of the regression of a given stock is lower than 0.01, this is accounted for in the corresponding column for that level, and that risk factor. The identical operation is performed for each stock and for each of the coefficients.

***T-STATISTIC* RESULTS**

	Significance Level							
Risk Factors	0.01	%	0.05	%	0.10	%	*Total to 0.10*	%
Intercept	12	*2*	31	*4*	51	*7*	94	*14*
RMRF	471	*68*	20	*3*	27	*4*	518	*75*
HML	102	*15*	79	*11*	50	*7*	231	*33*
SMB	99	*14*	90	*13*	71	*10*	260	*38*

TABLE 1

Notes: Percentages are calculated on the total securities considered, i.e. 692.

As can be seen, the market risk factor RMRF appears the greatest number of times as being statistically significant. Taking the 1% level, this factor is significant in 471 securities. Moreover, of the remainder, 20 appear significant at the 5% level and 27 more at the 10% level. Therefore, for a total of 518 stocks, i.e. 75% of the sample, the market risk factor RMRF is significant at the 10% level. Levy and Roll (2012) argue that the CAPM, in spite of the limitations it presents (because it only considers the market risk factor), is a useful model for estimating stock returns.

As for the SMB and HML risk factors, both appear as significant in a similar percentage at the different significance levels. Specifically, at the 1% level, the HML risk factor appears as significant for 102 stocks, as is the SMB risk factor for 99 securities. In addition, at the 5% level, 79 stocks appear significant for the HML and 90 for the SMB. Lastly, at the 10% level, 50 stocks more emerge for the HML and 71 for the SMB, giving us significance for a total of 231 stocks for the HML and 260 for the SML.

Once these results have been obtained, the analysis should be continued to obtain a robust conclusion about the ability of the model as a portfolio management tool, especially when taking into account the results with respect to the intercept.

The intercept only appears significant in 14% of the regressions and when the 10% significance level is taken, indicating that most of the stocks' behavior is captured by the risk factors considered in the model. A good model specification produces intercepts that are indistinguishable from zero (Merton, 1973). As Fama and French (1992) indicate, the intercept estimation provides a simple measure, and in turn a formal test, of how the different factors capture the average performance.

With the purpose of analyzing the ability of the model as a whole, the *F-statistic* is calculated. The results obtained are presented in Table 2:

***F-STATISTIC* RESULTS**

Significance Level	No. of Securities	Accumulated	%
1%	670	670	96.82
5%	13	683	98.70
10%	4	687	99.28

TABLE 2

Notes: Percentages are calculated on the total securities considered, i.e. 692.

Our results indicate that the model capacity as a whole, following the procedure developed for the construction of the factors, is very satisfactory. The model appears as statistically significant at the 1% level for 670 securities, i.e. for almost 97% of the sample. Taking the 5% level, 13 more stocks appear as significant, and if we move to the 10% level, the model is significant 99.28% of the time.

Given the results of this second test, it can be seen that the F&F model can, indeed, be adapted to variations in the characteristics of companies.

However, can the inferences provided by the model be taken as correct?

3.2. Econometric Analysis

In this section, the results of the several tests to check the assumptions of the coefficients estimation methodology are presented.

The errors have zero mean: $E(\mu_t) = 0$

This requirement is satisfied since the regression calculated includes a constant term.

The variance of the errors is constant and finite over all values of x_t: var (μ_t) = $\sigma^2 < \infty$

Table 3 shows the results obtained from the application of the White (1980) test to measure the heteroskedasticity on the errors.

WHITE TEST RESULTS

Significance Level	No. of Securities	Accumulated	%
10%	14	14	2.02
5%	19	33	4.77
1%	44	77	11.13

TABLE 3

Notes: Percentages are calculated on the total securities considered, i.e. 692.

The table displays the number of times the *p-value* associated with the White (1980) test in each of the regressions run for each stock appears as significant at the 1, 5 and 10% levels.

The results indicate that, for the 10% level, the null hypothesis of homoskedasticity should be rejected for a total of 14 stocks, i.e. the error of these securities does not present constant variance. At the 5% level, the null hypothesis should be rejected for 19 more stocks, and if it is taken to the 1% level, 44 stocks more, i.e. 77 of the 692 stocks considered for the study, 11.13%, present a heteroskedasticity problem in the errors.

This check, and the results obtained, highlights the need to exercise caution when such an analysis is carried out.

The consequences of using OLS estimators in the presence of heteroskedasticity is that, in spite of the fact that the coefficients are not biased and are consistent, these are not the best possible coefficients, i.e. they are not those that present the fewest standard errors. Therefore, for securities presenting such problems, incorrect estimates are performed.

Nevertheless, there are several ways to deal with heteroskedasticity, such as the logarithmic transformation and the standard errors correction, following White (1980), but, as indicated previously, our aim is to highlight the need to research this aspect before performing the analysis.

There is no relationship between the error and corresponding x variable: cov (μ_t, x_t) = 0

Table 4 presents the results of the Breusch-Godfrey test to measure the autocorrelation of the errors.

At the 10% level, the null hypothesis of no autocorrelation should be rejected for a total of 41 stocks, at the 5% it level should be rejected for 41 securities more, and another 12 at the 1% level, for a total of 94 stocks, 13.58% of the total, presenting autocorrelation on the residues.

BREUSCH-GODFREY TEST RESULTS			
Significance Level	No. of Securities	Accumulated	%
10%	41	41	5.92
5%	41	82	11.85
1%	12	94	13.58

TABLE 4

Notes: Percentages are calculated on the total securities considered, i.e. 692.

These results show that problems of autocorrelation exist with some stocks, and, just as with the heteroskedasticity problem, the assumption of no autocorrelation should be present. Despite the fact that the estimated coefficients, when there is autocorrelation, are not biased, these are inefficient and consequently incorrect estimates are obtained.

In the autocorrelation case, there are also different ways to deal with it, such as the Cochrane-Orcutt (1949) procedure, or the variance-covariance matrix of Newey and West (1987). In any case, the need to research this aspect is highlighted.

In this respect, authors such as Hendry and Mizon (1978) suggest that the presence of autocorrelation presents an opportunity rather than a problem, maintaining that serial correlation in the errors arises as a consequence of a mis-specified dynamic. (The dynamic models arise from this approach.)

There is no relationship between the error and the corresponding x variable: cov **$(\mu_t, x_t) = 0$**

As for the analysis of the correlation between the explanatory variables and the disturbance term for each of the regressions, the results obtained are:

CORRELATION LEVELS BETWEEN THE RISK FACTOR RMRF AND THE DISTURBANCE TERM			
Interval	No. of Securities	Interval	No. of Securities
[1;0.9)	0	[0;-0.1)	335
[0.9;0.8)	0	[-0.1;-0.2)	2
[0.8;0.7)	0	[-0.2;-0.3)	5
[0.7;0.6)	0	[-0.3;-0.4)	1
[0.6;0.5)	0	[-0.4;-0.5)	0
[0.5;0.4)	0	[-0.5;-0.6)	0
[0.4;0.3)	0	[-0.6;-0.7)	0
[0.3;0.2)	0	[-0.7;-0.8)	0
[0.2;0.1)	6	[-0.8;-0.9)	0
[0.1;0)	343	[-0.9;-1]	0
[0.01;-0.01] — > 630			

TABLE 5

Table 5 presents different correlation level intervals between the risk factor RMRF and the disturbance term. It is possible to appreciate that for a total of 343 stocks, the risk factor RMRF is positively correlated with the error term between 0 and 10%. If the interval between 0 and -10% is taken, for 335 stocks the RMRF is negatively correlated with the error term. Likewise, the table also shows the quantity of securities for which the risk factor is correlated with the error term in a percentage between -1 and 1%. Thus, for 630 stocks, the correlation level is less than ±1%.

These results point out that, for the overwhelming number of regressions run for each security, the correlation between the risk factor RMRF and the error term is very weak. Nevertheless, this result does not diminish the importance of the necessity of analyzing this aspect.

CORRELATION LEVELS BETWEEN THE RISK FACTOR SMB AND THE DISTURBANCE TERM

Interval	No. of Securities	Interval	No. of Securities
[1;0.9)	0	[0;-0.1)	336
[0.9;0.8)	0	[-0.1;-0.2)	15
[0.8;0.7)	0	[-0.2;-0.3)	5
[0.7;0.6)	0	[-0.3;-0.4)	0
[0.6;0.5)	0	[-0.4;-0.5)	0
[0.5;0.4)	0	[-0.5;-0.6)	1
[0.4;0.3)	0	[-0.6;-0.7)	0
[0.3;0.2)	1	[-0.7;-0.8)	0
[0.2;0.1)	14	[-0.8;-0.9)	0
[0.1;0)	320	[-0.9;-1]	0
[0.01;-0.01] — > 613			

TABLE 6

As can be seen in Table 6, the correlation between risk factor SMB and the disturbance term is within low levels. For a total of 613 stocks, the SMB factor presents correlations with the error term between -1 and 1%.

Nevertheless, this factor appears to be correlated with the error for a greater number of stocks between the 10 and 20% levels, both negative and positive. For a single stock, the correlated level is between -50 and -60%, concretely -0.5097. In spite of that, these correlation levels may be considered relatively low and acceptable, and therefore be ignored. Nevertheless, the process continues to be relevant, since greater correlations could exist.

CORRELATION LEVELS BETWEEN THE RISK FACTOR HML AND THE DISTURBANCE TERM			
Interval	No. of Securities	Interval	No. of Securities
[1;0.9)	0	[0;-0.1)	335
[0.9;0.8)	0	[-0.1;-0.2)	3
[0.8;0.7)	0	[-0.2;-0.3)	1
[0.7;0.6)	1	[-0.3;-0.4)	0
[0.6;0.5)	0	[-0.4;-0.5)	1
[0.5;0.4)	0	[-0.5;-0.6)	0
[0.4;0.3)	0	[-0.6;-0.7)	0
[0.3;0.2)	7	[-0.7;-0.8)	0
[0.2;0.1)	8	[-0.8;-0.9)	0
[0.1;0)	336	[-0.9;-1]	0
[0.01;-0.01] — > 621			

TABLE 7

Again in this case, as shown in Table 7, the number of the stocks for which the degree of correlation between the risk factor HML and error term comes within 1% interval is very large, 621. It is also clear that, for the remaining securities the correlation is not relevant, with the exception of one stock for which the correlation level is 0.6524. As a consequence, if the model is used on this stock, the estimation of the expected return will be inefficient since it is assigned to a power higher than the variable, when in reality this is due to the correlation between the error term and the explanatory variable.

Therefore, the results of the correlation levels between the risk factors and the disturbance term can be considered acceptable for most the securities, with the exception of the last case, so the adapted F&F model maintains the no-correlation hypothesis. On the other hand, this analysis highlights the need to research the compliance with this assumption, since the estimated parameters may be biased or inconsistent, and inefficient estimates will result, with important repercussions if they are accepted by portfolio managers.

The disturbances are normally distributed: $\mu_t \sim N(0, \sigma^2)$

Appealing to the central limit theorem, this has obviated the residual normality test. As mentioned previously, the normality absence of the distribution of the residuals is not sufficient condition to exclude the model's validity. The statistic tests follow a correct distribution, even when the residuals are not normal.

4. CONCLUSIONS

Given the results obtained from the significance analysis of the model as a whole, measured by the *F-statistic*, it is understood that the F&F model can be adapted and become a more versatile and flexible tool, capable of incorporating the variations in the characteristics of companies more consistently.

As for the results achieved concerning the explanatory power of the model's individual coefficients, it is possible to conclude that, in the first place, the limited number of times the intercept appears as significant provides positive evidence that most of the stock returns are being fitted by the risk factors considered in the model. In the second place, the procedure developed for the SMB and HML risk factors can be more flexible, with results that are somewhat more satisfactory in the SMB case, although both display significance in similar percentages. Finally, the RMRF factor is the one that appears as significant the greatest number of times.

Regarding the degree of maintaining the estimation methodology assumptions, the proposed model meets the desirable properties (BLUE) for the overwhelming number of securities. Consequently, it can be considered that, except for certain stocks to which several methodologies could be applied to solve specific problems, the inferences run from the model can be taken as being correct.

In this paper, we have highlighted the need to analyze that the hypotheses of the estimation methodology are maintained when inferences with a model such as that of Fama and French (1992) are performed, since otherwise, the estimates obtained will be incorrect, and portfolio managers will be faced with unexpected results.

REFERENCES

Abhakorn, P., Smith, P. N., and Wickens, M. R., 2013. What do the Fama–French factors add to C-CAPM? *Journal of Empirical Finance,* 22, 113-127.

Agarwal, V., and Poshakwale, S., 2010. Size and book-to-market anomalies and omitted leverage risk. *The European Journal of Finance,* 16 (3), 263-279.

Avramov, D., and Chordia, T., 2006. Asset pricing models and financial market anomalies. *Review of Financial Studies,* 19, 1001-1040.

Bartholdy, J., and Peare, P., 2005. Estimation of expected return: CAPM vs. Fama and French. *International Review of Financial Analysis,* 14, 407-427.

Black, F., 1972. Capital market equilibrium with restricted borrowing. *Journal of Business,* 45 (3), 444–455.

Brooks, C., 2014. *Introductory econometrics for finance*, 3rd ed. Cambridge University Press.

Cochrane, D., and Orcutt, G. H., 1949. Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association,* 44 (245), 32-61.

Durbin, J., and Watson, G. S., 1951. Testing for serial correlation in least squares regression. *Biometrika,* 38, 159-71.

Chen, D. H., Chen, C. D., and Chen, S. W., 2014. VaR and the crosssection of expected stock returns: an emerging market evidence. *Journal of Business Economics and Management,* 15 (3), 441-459.

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

Fama, E. F., and MacBeth, J., 1973. Risk, return and equilibrium: empirical tests. *Journal of Political Economy,* 81 (3), 607–636.

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

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

Fama, E. F., and French, K., 1996. The CAPM is wanted, dead or alive. *The Journal of Finance,* 51 (5), 1947–1958.

Ferruz, L., and Badía, G., 2015. A sector strategy from the Fama and French model. *Applied Economics Letters,* 22 (18), 1511–1514.

Fletcher, J., 2001. An examination of alternative factor models in UK stock returns. *Review of Quantitative Finance and Accounting,* 16 (2), 117-130.

Goldfeld, S. M., and Quandt, R. E., 1965. Some tests for homoscedasticity. *Journal of the American statistical Association,* 60 (310), 539-547.

Grauer, R. R., and Janmaat, J. A., 2010. Cross-sectional tests of the CAPM and Fama–French three-factor model. *Journal of Banking and Finance,* 34, 457-470.

Gregory, A., Tharyan, R., and Christidis, A., 2013. Constructing and testing alternative versions of the Fama–French and Carhart models in the UK. *Journal of Business Finance and Accounting,* 40 (1-2), 172-214.

Gujarati, D. N., 2003. *Basic Econometrics*. 4th ed. McGraw-Hill, New York.

Hamard A. and Mascareñas J., (2010). La Prima de Riesgo por Tamaño en el Mercado Continuo Español. *Análisis Financiero,* 114, 34-40.

Hendry, D. F., and Mizon, G. E., 1978. Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England. *The Economic Journal,* 88, 549-563.

Hill, R. C., Griffiths, W. E., and Judge, G. G., 1997. *Undergraduate Econometrics.* Wiley, New York

Hussain, S. I., Toms, J. S., and Diacon, S., 2002. Financial distress, market anomalies and single and multifactor asset pricing models: New evidence. Available from Internet: http://ssrn.com/abstract=313001

Hwang, S., and Satchell, S. E., 2005. Valuing information using utility functions: how much should we pay for linear factor models? *The European Journal of Finance,* 11 (1), 1-16.

Jensen, M. C., 1968. The performance of mutual funds in the period 1945-1964. *The Journal of Finance,* 23 (2), 389-416.

Lajili-Jarjir, S., 2007. Explaining the Cross-section of Stock Returns in France: Characteristics or Risk Factors? *The European Journal of Finance,* 13 (2), 145-158.

Levy, M., and Roll, R., 2012. A New Perspective on the Validity of the CAPM: Still Alive and Well. *Journal of Investment Management,* 10 (3), 9-20.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics,* 47 (1), 13–37.

Merton, R. C., 1973. An intertemporal capital asset pricing model. *Econometrica,* 41, 867-887.

Miles, D., and Timmermann, A., 1996. Variation in expected stock returns: evidence on the pricing of equities from a cross-section of UK companies. *Econometrica,* 63 (251), 369-382.

Newey, W. K., and West, K. D., 1987. A simple positive semi-definite heteroskedasticity and autocorrelation-consistent covariance matrix. *Econometrica,* 55, 703-8.

Nichol, E., and Dowling, M., 2014. Profitability and Investment Factors for UK Asset Pricing Models. *Economics Letters,* 125 (3) 364-366.

Pettengill, G., Chang, G., and Hueng, J., 2012. Risk-return Predictions with the Fama-french Three-factor Model Betas. *International Journal of Economics and Finance,* 5 (1), 34.

Rogers, P., and Securato, J. R., 2007. Comparative Study of CAPM, Fama and French Model and Reward Beta Approach in the Brazilian Market. Available from Internet: http://ssrn.com/abstract=1027134

Sharpe, W. F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance,* 19 (3), 425–442.

Soumaré, I., Aménounvé, E. K., Diop, O., Méité, D., and N'sougan, Y. D., 2013. Applying the CAPM and the Fama–French models to the BRVM stock market. *Applied Financial Economics,* 23, 275-285.

Treynor, J., 1965. How to rate management of investment funds. *Harvard Business Review,* 43 (1), 63-75.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica,* 48, 817-838.

Zhong, A., Limkriangkrai, M., and Gray, P., 2014. Anomalies, risk adjustment and seasonality: Australian evidence. *International Review of Financial Analysis,* 35, 207-218.