Economics and Business Letters
14(2), 75-87, 2025

Factors relevance in asset pricing: new evidences in emerging markets from random matrix theory

Laura Molero-González[1*] • Juan E. Trinidad-Segovia[2] • Miguel A. Sánchez-Granero[3] • Andrés García-Medina[4]

[1] *Department of Economics and Business, University of Almería, Almería, Spain & Department of Social Sciences and Economics, Sapienza University of Rome, Rome, Italy*
[2] *Department of Economics and Business, University of Almería, Almería, Spain*
[3] *Department of Mathematics, University of Almería, Almería, Spain*
[4] *Faculty of Science, Autonomous University of Baja California (UABC), Ensenada, Mexico*

Received: 25 May 2024
Revised: 27 September 2024
Accepted: 8 November 2024

Abstract
In financial literature, there is an ongoing debate as to whether multi-factor models provide better results in explaining the cross-sectional expected return of financial assets than the Sharpe Simple Index model. Despite the evidence provided by some authors about the superiority of market Beta in major developed markets, the debate does not seem to be closed, even less so for emerging markets. In this paper, we provide new evidence on the number of significant factors in emerging markets using Random Matrix Theory (RMT) statistical techniques. We find that, with a confidence level of 99%, no significant factors are found in emerging markets, compared to developed ones, where market beta is always the unique factor. Our results confirm that emerging markets have different characteristics in relation to developed markets, which is evident in the factors affecting the performance of financial assets.

Keywords: emerging markets, asset pricing, random matrix theory, single index model, APT
JEL Classification Codes: G12, G15, C53, D53

1. Introduction

[*] Corresponding author. E-mail: lmg172@ual.es.

Citation: Molero-González, L., J.E. Trinidad-Segovia, M.A. Sánchez-Granero and A. García-Medina (2025) Factors relevance in asset pricing: new evidences in emerging markets from random matrix theory, *Economics and Business Letters*, 14(2), 75-87. DOI: 10.17811/ebl.14.2.2025.75-87.

Random matrix theory (RMT) can be thought of as a new type of statistical mechanics where instead of having a set of states governed by the same Hamiltonian, we have a set of Hamiltonians governed by the same symmetry. Historically, this theory was introduced into mathematical statistics by Wishart (1928).

Many mathematicians later worked out of purely theoretical interest. The physical theory foundations of RMT can be traced back to the 1950s decade when Wigner proposed a statistical description of the energy levels of the uranium nucleus (Mehta, 2004) employing RMT. In 1962, Dyson extended Wigner's ideas, showing that physically reasonable symmetry assumptions can be represented by Gaussian ensembles (Dyson, 1962).

This theory gained momentum when (Bohigas et al., 1984) stated the quantum chaos conjecture.

We owe its connection with finance, and especially with portfolio theory, to the seminal works of Laloux et al. (1999); Plerou et al. (2002) where the authors propose the use of RMT to model the interactions of financial markets through the Wishart ensemble. The general implications of this ensemble are framed within multivariate statistics. Our interest in this work is to go further and extend the applicability of RMT to the particular topic of asset pricing for emerging markets.

There exist several proposals to model asset pricing. Together with the mean-variance Markowitz model, the introduction of the Single Index Model (SIM) by Sharpe (1964) is one of the most relevant contributions made to the study of financial markets. The implications of this model have been important not only for researchers, but also for practitioners, to the extent that it is associated with the performance of a financial asset with its systematic risk, which is measured by Beta. Mathematically, the model is expressed as follows:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad (1)$$

where R_{it} is the return on the security i for the period t, Rmt is the market return (usually proxies through its main equity index or a value-weighted portfolio of all shares), and ε_{it} is a zero-mean residual.

The first empirical tests of this model were performed by F. Black and Scholes (1972) and Fama and MacBeth (1973), who found that it was valid to explain the returns of companies before 1969. However, lately Reinganum (1981), Lakonishok and Shapiro (1986) and Fama and French (1992) found anomalies for different periods untił1990. The most obvious reason seemed to be the existence of additional factors relevant to asset pricing. The most important contribution in this line is the Fama and French Three-Factor Model (Fama and French, 1992, 1995), mathematically expressed as:

$$R_{it} - R_f = \alpha_i + \beta_i\left(R_{mt} - R_f\right) + s_i SMB_t + h_i HML_t + \varepsilon_{it} \quad (2)$$

where SMB_t is the return on a diversified portfolio of small stocks minus the returns on a diversified portfolio of large stocks, and HML_t is the difference between the returns on diversified portfolios of high book-to-market and low book-to-market stocks.

Unfortunately, the three-factor model fails in explaining the cross-sectional expected returns for extreme growth and microcap extreme growth stocks. Consequently, in 2015 the authors propose the Five-Factor Model (Fama and French, 2015), which is defined as:

$$R_{it} - R_f = \alpha_i + \beta_i(R_{mt} - R_f) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \varepsilon_{it} \quad (3)$$

where RMW_t is the difference between the returns in diversified portfolios of stocks with robust and weak profitability correspondingly, and CMA_t is the difference between the returns in diversified portfolios of stocks of companies with low and high investment practices.

Since the empirical tests of Arbitrage Pricing Theory (APT) models suggested that some anomalies persisted (see, for example, Tilman et al. (2004), Novy-Marx (2013), Fama and French (2017) or Kubota and Takehara (2018)), researchers have proposed a wide number of different factors. Harvey et al. (2016) summarizes 313 papers that study cross-sectional return patterns providing a taxonomy of historical factors and definitions.

A classic debate is one that seeks to know whether the single-factor model is dead or not considering all this "zoo factor". First, Isakov (1999), later Racicot and Rentz (2016) and De Nard et al. (2021), and recently Molero-González et al. (2023) concluded that the Sharpe Simple Index Model explains the expected cross-sectional return better than the APT models.

However, if this debate remains open in developed markets, emerging markets will be even better. In the financial literature, it is clear that emerging markets show very different patterns of behavior than developed markets: higher variability, higher serial correlation, and informational inefficiency (Buckberg, 1995; Balladares et al., 2021). Assets Pricing in emerging markets is not a fully developed topic. The pioneering paper of Buckberg (1995) found that for 1977-1984, in 60% of the emerging markets studied, the SIM did not provide a consistent and efficient estimation of Beta. In the period 1985-1991, a high rate-of-return variance impedes the estimation of market betas.

Garcia and Ghysels (1998) tested two Capital Asset Pricing Models (CAPM): a conditional world CAPM and a simple local CAPM model for a set of emerging markets. The authors concluded that for the conditional world CAPM the model yields a system- atic mispricing of risk factors. The CAPM model for size-ranked portfolios showed a much more stable relationship than a simple local CAPM model.

A recent part of the financial literature has explored the factors anomalies in emerging markets. Cakici et al. (2013) studied value and momentum in 18 emerging stock markets, finding strong evidence for the effect of value in all emerging markets and the effect of momentum for all but Eastern Europe.

Cakici et al. (2016) analyzed the effects of size, value, and momentum in 18 emerging stock markets during the 1990-2013 period, finding that only the value effect exists in all markets except Brazil.

Lalwani and Chakraborty (2020) study the performance of various multifactor asset pricing models in ten emerging and developed markets, finding that the parsimonious three-factor model or its four-factor variants appear to be more suitable across all markets.

In this paper we pretend to provide new evidence on the significant factors that explain the cross sectional expected return of stocks in emerging markets and we do it in the statistical sense of high dimensionality. Unlike other works previously published in the financial literature, the methodology proposed allows to assume other distributions beyond the Gaussian, even with autocorrelations, under the unique condition of having a finite fourth moment. On the other hand, It is a purely statistical and does not need regularization techniques or heuristic strategies.

The paper is structured as follows: Section 2 introduces the RMT statistical tests, on which this study is based; Section 3 brings up the methodology; Section 4 presents the data and shows and discusses the results obtained; and Section 5 concludes.

2. Random matrix theory statistical tests

When discussing high-dimensionality in statistics, it is well known that the covariance matrix is distributed according to the Wishart distribution (Wishart, 1928). Consider X a matrix of dimension n p with random entries identically and independently distributed (IID) according to a Gaussian probability distribution function (PDF) with mean zero and variance equal one. Each vector X_i of dimension $p\,x\,n$ is denoted as $X_i \sim N_p(0, \Sigma), i = 1, \ldots, n$. In statistics $W = X^T X$ is said to have a p-variant Wishart distribution of n degrees of freedom $W_p\,(n, \Sigma)$.

In RMT we say that the matrices with structure W belong to the Wishart Orthogonal Ensemble (WOE), which are rotationally invariant matrices under the Haar measure. A universal result is that regardless of the particular matrix X , when the number of columns p is of the same order as the number of rows n, and both dimensions grow without any limit, then the distribution of eigenvalues of the matrix W converges to what is known as the Marchenko-Pastur law (Marchenko and Pastur, 1967).

$$\rho(\lambda) = \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi q \lambda}, \lambda_\pm = (1 \pm \sqrt{q})^2 \quad (4)$$

where is assumed that $\Sigma = I$.

In the financial context, X represents the data matrix of returns, where the first dimension usually represents the number of transaction days, and the second dimension the number of assets. Thus, W becomes the covariance matrix of p return time series of length n. Hence, the Marchenko-Pastur law implies a bias in the estimation of the covariance matrix that increases as $q = {}^{p}/_{n} \rightarrow \infty$ when $n, p \rightarrow \infty$.

On the other hand, the asset pricing problem relies on the factor models, which in turn have the covariance matrix as their central object. Consequently, we need to take into account the effect of the dimension to avoid bias in estimating the number of factors that determine the price of the portfolio assets.

To address this statistical bias, one approach is to consider the Tracy-Widom distribution (Tracy and Widom, 1994), which elucidates the behavior of eigenvalues near the boundary of the Marchenko-Pastur law. This distribution helps discern whether eigenvalues reside within the noise zone or signify true signals. By leveraging the Tracy-Widom distribution, hypothesis tests can be constructed to formally differentiate this bias.

Onatski (2008) extended this work, demonstrating that the distribution of the first r centered and scaled eigenvalues of the Wishart matrix with complex entries converges weakly to the r-dimensional joint distribution of the Tracy-Widom distribution, denoted as F_2.

In this way, Onatski (2009) proposed the R statistic to statistically determine the number of factors in the context of the generalized dynamic factor model (DFM) (Forni et al., 2000). Below we show the methodology to apply the R statistic in the simplified problem when the loading matrix does not depend on the lag operator. Therefore, we can test the hypothesis of the approximate (Chamberlain and Rothschild, 1982) (in contrast to the dynamic) number of factors.

3. Materials and methods

Consider the data matrix X composed of the price returns of p firms over n trading days. The first step is to divide X into two submatrices X_1 and X_2 of dimension $\frac{n}{2}p$, that is, X_1 contains the returns at times $t = 1, \dots, \frac{n}{2}$, and the matrix X_2 in the consecutive period $t = \frac{n}{2} + 1, \dots, n.$ The complex matrix $\hat{X}$ is then constructed as:

$$\hat{X} = X_1 + iX_2 \tag{5}$$

The second step consists of calculating the eigenvalues $\hat{\lambda}_1, \dots, \hat{\lambda}_p$ of the covariance matrix $\frac{2}{n}\hat{X}\hat{X}^{\ddagger}$[1].

In a third step, the statistic $\hat{R}$ is calculated as a function of the eigenvalues obtained in the previous step:

$$\hat{R} = \max_{k_0<i<k_1} \frac{\hat{\lambda}_i - \hat{\lambda}_{i+1}}{\hat{\lambda}_{i+1} + \hat{\lambda}_{i+2}} \tag{4}$$

The fourth step consists of assuming that the real number of factors k is between k_1 and k_2. Then construct the null hypothesis $H_0: k = 1$ versus $H_1: k_1 < k \leq k_2$ at the α level of significance using the critical values of the statistic $\hat{R}$ to accept or reject the null hypothesis[2]. If H_0 is accepted, the number of statistically significant factors is k_1 and the algorithm is stopped. If, in contrast, H_0 is rejected, we try $H_0: k = k_1 + 1$ versus $H_1: k_1 + 1 < k \leq k_2$. The procedure is repeated until H_0 is not rejected and the corresponding number of statistically significant factors is considered the best estimate of the associated model.

The general idea of the algorithm is to contrast the value of the R statistic with respect to the critical values iteratively for a different number of factors. This is repeated until there is no evidence to reject the null hypothesis. In this way, it is possible to statistically confirm the validity of the models that are distinguished by proposing a different number of factors to determine the asset prices.

Regarding data, two groups of datasets were considered for this study: developed markets and emerging markets. Seven developed markets have been studied: Spain, France, Germany, London, Italy, Japan, and United States. Ten emerging markets were considered: China, Brazil, Argentina, Chile, India, Korea, Greece, Taiwan, Turkey and Malaysia. This classification of countries is based on that made by MSCI Inc.

MSCI Inc. evaluates equity markets around the world each year and classifies them into four categories: developed, emerging, frontier, and standalone markets. This classification is made

[1] Symbol ‡ represents the transpose of conjugate of the matrix.
[2] See Appendix B.

according to three criteria: market accessibility, economic development, size, and liquidity. For further reading, see MSCI (2023).

A selection criterion was established by which only those companies with less than 5% of missing days were taken. Then the logarithmic returns were estimated. The study period runs from January 2005 to December 2022.

In the context of RMT, the determining parameter is the ratio between the dimensions $q = \frac{p}{n}$. In particular, the noise effect due to the finiteness of the sample is observed for values of $q \to 1$. Therefore, in this study, the values of $q = \frac{1}{2}$ were considered to analyze the explanatory power of the test. The scenario of low-dimensionality can be solved with classical statistics. However, the case $q = \frac{1}{2}$ of high-dimensionality requires modern techniques such as the Tracy-Widom and Onatski statistical tests, which come from RMT.

Specifically, a sliding window experiment was designed for each market. Then, the number of significant factors was determined in each window at $\alpha = \{0.01, 0.05, 0.1\}$. Additionally, an upper bound $k_2 = 8$ was considered, so the alternative hypothesis was tested only up to that limit, enough to cover most situations.

The number of observations (n) and companies (p) for each market and window are shown in Table 1. In both datasets, we consider a skip of Δn=20 days of transactions (roughly a month in calendar days) for successive windows.

Table 1. Summary of the number of companies and observations for each of the markets studied

Country	**Market**	**p**	**n**
Spain	IBEX35	21	42
France	CAC40	33	66
Germany	DAX40	29	58
United States	Dow Jones	28	56
England	FTSE 100	79	158
Italy	FTSE MIB	25	50
Japan	Nikkei	206	412
China	SSE50	13	26
Brazil	Ibovespa	23	46
Argentina	Merval	12	24
Chile	S&P CLX IPSA	17	34
India	BSE SENSEX	26	52
Korea	KOSPI	264	528
Greece	ATHEX	40	80
Taiwan	TWSE	31	62
Turkey	BIST30	19	38
Malaysia	KLCI	21	42

Source: own elaboration

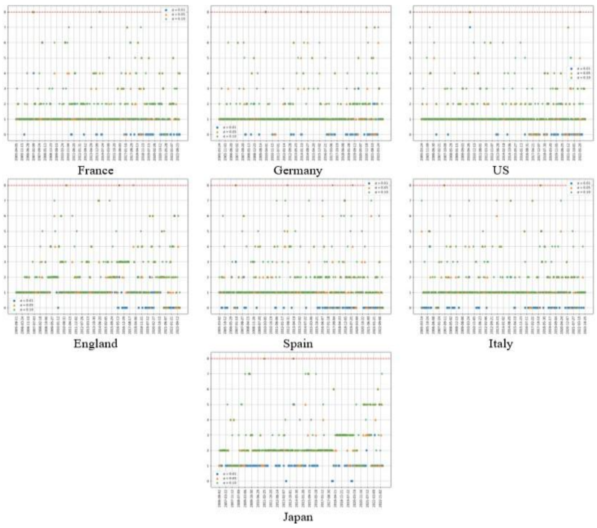
4. Results and discussion

Figure 1 shows the dynamic of the estimated number of factors for developed markets.

As can be seen, all of these markets have a common behavior. For a confidence level of 99%, there is only one statistically significant factor, with the exception of two periods (2018-2020 and 2021-2022), for which the test does not identify factors. If we reduce the confidence level to 95% and 90%, in general, there is only one factor, with the exception of small periods, where a two-factor model can be considered.

As in Molero-González et al. (2023), we have tested whether the identified factor corresponds to the Sharpe market factor. To do so, we repeated the experiment, applying the test to the covariance matrix of the residual of the Sharpe model (see equation 1), instead of applying it to the covariance matrix of the returns. In this way, when the market is removed, no factors are identified for the three alpha levels, with the exception of those periods for which two factors were identified. In these periods, now the test identifies only one factor, letting us confirm that the market is the sole explanatory factor identified by the test and one of the two-factor model. These results are in line with those obtained in Molero-González et al. (2023). The graphs corresponding to these results are shown in Appendix B.

Figure 1. Number of factors as a function of time for the group of developed markets.



Source: own elaboration
Note: The parameter α denotes the significance level of the Onatski test,
and the red dotted line represents the alternative hypothesis's upper bound k_2

An exception to this general behavior is Japan. For a confidence level of 99%, there is only one factor, with some periods where there are two. Unlike the other markets, we do not identify the two periods with no factors at the end of the study period. For the confidence levels of 95% and 90%, up to five factors are identified by the test. When the market factor is removed, the results are the same as for the rest of the markets: one less factor is identified in all cases.

Figure 2. Number of factors as a function of time for the group of emerging markets

Source: own elaboration
The parameter α denotes the significance level of the Onatski test,
and the red dotted line represents the alternative hypothesis's upper bound k_2.

Figure 2 shows the dynamic of the number of factors for the group of emerging markets studied. As happened with developed markets, here we also observe general behavior. Taking into account $\alpha = 0.01$, there are no significant factors. If the level α increases to 0.05 and 0.1, the test identifies only one factor as significant in explaining the cross-section of stock returns. If we remove the market factor, there are zero factors in all cases. The market seems to be less significant for these markets than for developed ones. The graphs presenting the results when the market factor is removed are shown in Appendix B.

Two exceptions to this general trend are Korea and Greece, which show a behavior more similar to that observed in Figure 1, being the market factor significant at a confidence level of 99% for these two markets. So, despite that Korea and Greece are considered emerging markets, they exhibit an analogous behavior to developed markets.

5. Conclusions

The financial literature on APT has proposed innumerable factors to explain market anomalies that were not captured by market Beta. A permanent debate has been whether models with more than one factor, such as those of Fama and French (1995, 2015), explained the cross-sectional of expected returns more effectively. Authors such as Isakov (1999), Racicot and Rentz (2016) or Molero-González et al. (2023) provided further evidence on the superiority of market Beta in major developed markets. Despite this, the debate does not seem to be closed for developed markets, but even less so for emerging markets.

Through a purely statistical approach, we have provided new evidence on the number of significant factors in explaining cross-sectional returns, differentiating between developed and emerging markets. Using an RMT statistical technique, at a confidence level of 99%, it is not possible to find any significant factor that explains the expected cross-sectional return of the considered emerging markets. When reducing the confidence level to 95% and 90%, the market factor seems to be the unique factor.

These results seem to confirm previous findings of Buckberg (1995) or Balladares et al. (2021) indicating that emerging markets have different characteristics than developed markets, which is also evident in the factors that affect the performance of financial assets.

However, it is important to remark that within the group of emerging markets, we have observed how two countries that are considered as emerging but turn out to behave as developed: it is the case of Korea and Greece.

MSCI considers Greece to be emerging because it does not meet the size and market accessibility requirements. In the case of Korea, the criteria that remain to be met are those related to market accessibility. Despite this, both markets perform similarly to developed markets. This fact leads us to think that perhaps the classification taken is not entirely accurate when considering countries as developed or emerging.

Acknowledgements. Laura Molero-González and Juan E. Trinidad-Segovia are supported by grant PID2021-127836NB-I00 (Ministerio Español de Ciencia e Innovación and FEDER). Miguel A. Sánchez-Granero acknowledges the support of grant PID2021-127836NB-I00 (Ministerio Español de Ciencia e Innovación and FEDER) and CDTIME.

Data availability. Data from the different markets are available on Yahoo Finance. Random Matrix Theory's codes are available at https://github.com/agarciam/RMT_statistical_tests/.

References

Balladares, K., Ramos-Requena, J.P., Trinidad-Segovia, J.E. and Sánchez-Granero, M.A. (2021) Statistical arbitrage in emerging markets: a global test of efficiency, *Mathematics*, 9(2), 179.

Black, F., Jensen, M. C. and Scholes, M. (1972) *The capital asset pricing model: some empirical tests*, in M.C. Jensen (ed.) Studies in the theory of capital markets, 79–121, Praeger.

Bohigas, O., Giannoni, M.-J. and Schmit, C. (1984) Characterization of chaotic quantum spectra and universality of level fluctuation laws, *Physical Review Letters*, 52(1), 1.

Buckberg, E. (1995) Emerging stock markets and international asset pricing, *The World Bank Economic Review*, 9(1), 51–74.

Cakici, N., Fabozzi, F.J. and Tan, S. (2013) Size, value, and momentum in emerging market stock returns, *Emerging Markets Review*, 16, 46–65.

Cakici, N., Tang, Y. and Yan, A. (2016) Do the size, value, and momentum factors drive stock returns in emerging markets? *Journal of International Money and Finance*, 69, 179–204.

Chamberlain, G. and Rothschild, M. (1982) Arbitrage, factor structure, and mean- variance analysis on large asset markets, *Econometrica*, 51(5), 1281-1304.

De Nard, G., Ledoit, O. and Wolf, M. (2021) Factor models for portfolio selection in large dimensions: The good, the better and the ugly, *Journal of Financial Econometrics*, 19(2), 236–257.

Dyson, F.J. (1962) The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics, *Journal of Mathematical Physics*, 3(6), 1199–1215.

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

Fama, E. and French, K. (1995) Size and book-to-market factors in earnings and return, *Journal of Finance*, 50(1), 131–155.

Fama, E. and French, K. (2015) A five factor asset pricing model, *Journal of Financial Economics*, 116, 1–22.

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

Fama, E.F. and French, K.R. (2017) International tests of a five-factor asset pricing model, *Journal of Financial Economics*, 123(3), 441–463.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000) The generalized dynamic-factor model: identification and estimation, *Review of Economics and Statistics*, 82(4), 540–554.

Garcia, R. and Ghysels, E. (1998) Structural change and asset pricing in emerging markets, *Journal of International Money and Finance*, 17(3), 455–473.

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

Isakov, D. (1999). Is beta still alive? Conclusive evidence from the Swiss stock market, *The European Journal of Finance*, 5(3), 202–212.

Jarrow, R.A. and Silva, F.B.G. (2015) Risk measures and the impact of asset price bubbles, *Journal of Risk,* 17(3), 35-56.

Kubota, K. and Takehara, H. (2018) Does the Fama and French five-factor model work well in Japan? *International Review of Finance*, 18(1), 137–146.

Lakonishok, J. and Shapiro, A. (1986) Systematic risk, total risk and size as determinants of stock market returns, *Journal of Banking and Finance*, 10(1), 115–132.

Laloux, L., Cizeau, P., Bouchaud, J.-P. and Potters, M. (1999) Noise dressing of financial correlation matrices, *Physical Review Letters*, 83(7), 1467.

Lalwani, V. and Chakraborty, M. (2020) Multi-factor asset pricing models in emerging and developed markets, *Managerial Finance*, 46(3), 360–380.

Marcenko, V.A. and Pastur, L.A. (1967) Distribution of eigenvalues for some sets of random matrices, *Mathematics of the USSR-Sbornik*, 1(4), 457.

Mehta, M.L. (2004) *Random matrices,* Elsevier.

Molero-González, L., Trinidad-Segovia, J., Sánchez-Granero, M. and García-Medina, (2023) Market beta is not dead: an approach from random matrix theory, *Finance Research Letters*, 103816.

MSCI (2023) Market classification framework.

Reinganum, M.R. (1981) Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values, *Journal of Financial Economics*, 9(1), 19– 46.

Novy-Marx, R. (2013) The other side of value: the gross profitability premium, *Journal of Financial Economics*, 108(4), 1–28.

Onatski, A. (2008) The Tracy–Widom limit for the largest eigenvalues of singular complex wishart matrices, *The Annals of Applied Probability*, 18(2), 470–490.

Onatski, A. (2009) Testing hypotheses about the number of factors in large factor models, *Econometrica*, 77(5), 1447–1479.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., Guhr, T. and Stanley, H.E. (2002) Random matrix approach to cross correlations in financial data, *Physical Review E*, 65(6), 066126.

Racicot, F.-E. and Rentz, W.F. (2016) Testing Fama–French's new five-factor asset pricing model: evidence from robust instruments, *Applied Economics Letters*, 23(6), 444–448.

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

Titman, S., Wei, K.J. and Xie, F. (2004) Capital investments and stock returns, *Journal of Financial and Quantitative Analysis*, 39(4), 677-700.

Tracy, C.A. and Widom, H. (1994) Level-spacing distributions and the airy kernel, *Communications in Mathematical Physics*, 159(1), 151–174.

Wishart, J. (1928) The generalised product moment distribution in samples from a normal multivariate population, *Biometrika*, 32–52.

Appendix A. Onatski Test

Table 2. Critical values for the $\hat{R}$ statistic

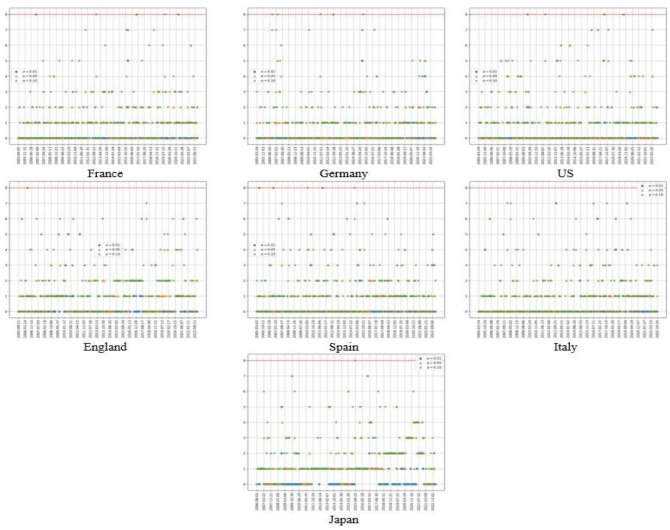
Size / %	$k_1 - k_0$							
	1	2	3	4	5	6	7	8
15	2.75	3.62	4.15	4.54	4.89	5.2	5.45	5.7
10	3.33	4.31	4.91	5.4	5.77	6.13	6.42	6.66
9	3.5	4.49	5.13	5.62	6.03	6.39	6.67	6.92
8	3.69	4.72	5.37	5.91	6.31	6.68	6.95	7.25
7	3.92	4.99	5.66	6.24	6.62	7	7.32	7.59
6	4.2	5.31	6.03	6.57	7	7.41	7.74	8.04
5	4.52	5.73	6.46	7.01	7.5	7.95	8.29	8.59
4	5.02	6.26	6.97	7.63	8.16	8.61	9.06	9.36
3	5.62	6.91	7.79	8.48	9.06	9.64	10.11	10.44
2	6.55	8.15	9.06	9.93	10.47	11.27	11.75	12.13
1	8.74	10.52	11.67	12.56	13.42	14.26	14.88	15.25

Source: Onatski (2009)
Note: The rows represent the level of significance and the columns the size of the test

Appendix B. Testing whether the market factor is the only identified factor

Figures 3 and 4 show the dynamic of the estimated number of factors when the market one is removed, for the group of developed markets (B1) and for emerging markets (B2). In other words, it shows the number of statistically significant factors identified by the Onatski test when the test is applied to the covariance matrix of the residual instead of applying it to the covariance matrix of returns.

Figure 3. Number of factors as a function of time for the group of developed markets, when the market factor is removed



Source: own elaboration
The parameter α denotes the significance level of the Onatski test,
and the red dotted line represents the alternative hypothesis's upper bound k_2.

Figure 4. Number of factors as a function of time for the group of emerging markets, when the market factor is removed

Source: own elaboration

Note: The parameter α denotes the significance level of the Onatski test, and the red dotted line represents the alternative hypothesis's upper bound k_2.