

# Extraction of the Underlying Structure of Systematic Risk from Non-Gaussian Multivariate Financial Time Series Using Independent Component Analysis: Evidence from the Mexican Stock Exchange

Rogelio Ladrón de Guevara Cortés<sup>1</sup>, Salvador Torra Porras<sup>2</sup>, Enric Monte Moreno<sup>3</sup>

<sup>1</sup> Veracruzana University,  
Institute for Research and Graduate Studies in Administrative Sciences (IIESCA),  
Mexico

<sup>2</sup> University of Barcelona, Faculty of Economics and Business,  
Department of Econometrics, Statistics and Applied Economy,  
Spain

<sup>3</sup> Polytechnic University of Catalonia,  
Barcelona School of Telecommunications Engineering,  
Department of Signal Theory and Communications,  
Spain

roladron@uv.mx, storra@ub.edu, enric.monte@upc.edu

**Abstract.** Regarding the problems related to multivariate non-Gaussianity of financial time series, i.e., unreliable results in extraction of underlying risk factors -via Principal Component Analysis or Factor Analysis-, we use Independent Component Analysis (ICA) to estimate the pervasive risk factors that explain the returns on stocks in the Mexican Stock Exchange. The extracted systematic risk factors are considered within a statistical definition of the Arbitrage Pricing Theory (APT), which is tested by means of a two-stage econometric methodology. Using the extracted factors, we find evidence of a suitable estimation via ICA and some results in favor of the APT.

**Keywords.** Extraction techniques, underlying risk factors, independent component analysis, arbitrage pricing theory, Mexican stock exchange.

## 1 Introduction

The goal of the present paper is to determine the statistical pervasive systematic risk factors in the Mexican Stock Exchange by means of an uncommon computational technique, namely, Independent Component Analysis (ICA), in order to detect a more reliable structure of the pervasive factors driving the returns on equities in the

Mexican Stock Exchange (BMV for its acronym in Spanish).

Because of its nature, ICA is designed by assuming a linear mixture of random variables that are not normally distributed, which is a relevant property for the problem we are dealing with. This technique helps to reveal a linear combination of underlying time series; by extracting their statistically independent components, the pervasive sources of some observed parallel time series can be explained.

ICA has been used, mainly in fields such as signal and image processing, speech and audio separation, biomedical signals and image analysis, telecommunications, neurophysiology, text and document processing, bioinformatics, environmental issues and some industrial applications. In relatively recent years, studies about the applications of ICA in different fields of Finance have been made in some countries.

The works that we considered more relevant in the context of our research have used ICA for extracting the following: the underlying factors explaining the stock returns in Japan [2], Hong

Kong [4], Italy [9], the USA [24] and during the crisis period [25]; the relevant factors driving the movements from implied volatility surfaces of index options [1]; the factors driving the movements of a term structure on interest rates in Germany [35]; the factors driving spot rate curve movements in the USA [3]; the factors moving the returns for real estate investment trusts in the USA [30], and for estimating the factor model of returns for the USA Thrift Saving Plan Funds [37], and the factors for pricing multiasset derivatives [26].

Moreover, some other representative studies of ICA in Finance have used this technique for the following purposes:

- (1) to analyze the interactions between currencies in the Foreign Exchange [36];
- (2) to model the conditional higher moments risk in international stock markets [48], the term structure of multiple yield curves [46], and the volatility of market price indexes [47];
- (3) to manage investment portfolios [8];
- (4) to allocate assets [32];
- (5) to forecast financial time series [30];
- (6) to compute improved portfolio risk measures such as VaR in banking sector [6, 7];
- (7) to explain the volatility of investment funds [45];
- (8) to generate an equity sector classification [43];
- (9) to improve bank performance evaluation [29];
- (10) to produce multifactor index variance from the SPX sector ETF returns [38];
- (11) to measure the dependency between stocks in the USA [17], and
- (12) to analyze herding among hedge fund styles [27].

As far as we are concerned, there is no study regarding the application of the ICA in Finance focused on Mexico. Consequently, we shall try to fill this gap in financial literature by contributing with the application of a novel extraction technique to extract the underlying structure of risk factors in the Mexican Stock Exchange.

The outline of this paper is as follows. In section 2, we briefly describe the ICA technique; in section

3, we present an empirical study; and in section 4, we draw the main conclusions.

## 2 Independent Components Analysis

### 2.1 ICA Basics

Despite the widespread evidence concerning the non-Gaussianity of the returns on equities, the most popular latent variables analysis techniques used for extracting the pervasive factors underlying the financial multivariate data are Principal Component Analysis (PCA) and Factor Analysis (FA), which assume a Gaussian distribution of the latent factors.

ICA represents an improved extraction technique for this kind of data, since it is based on a multivariate non-normality approach and looks for mutually and statistically independent components. According to [21], statistical independence means that not one of the components gives any information about the others.

Also following [10], mutually and statistically independent can be interpreted as being of different nature. ICA was introduced in the field of signal processing and neural computation as a tool to solve the problem of Blind Source Separation (BSS) and Signal Reconstruction.

According to [40], the former concept implies revealing hidden factors from observable measures, where we know very little about the original signals and their process of generation.<sup>1</sup> The basic technique for solving this kind of problem is ICA, which assumes that the observed variables are the result of an unknown mixing process of some latent original sources. Consequently, the observed variables can be decomposed by means of a demixing process, capable of estimating some statistically independent components that can be considered as reliable proxies for the original sources that generated the observed variables ( $s \approx y$ ).

The main characteristic of the latent sources is that they are assumed to be non-Gaussian and mutually independent. They are known as the

<sup>1</sup> According to [44] there are two approaches to solve the BSS problem: one based on the Independent Component Analysis and another based on Second Order Statistics.

independent components of the multivariate observed data.

According to [5], the formal expressions of the mixing and demixing processes in the basic ICA model are as follows:

$$\text{Mixing process: } \mathbf{x} = \mathbf{A}\mathbf{s}, \quad (1)$$

$$\text{Demixing process: } \mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s}. \quad (2)$$

where  $\mathbf{x}$  represents the vector of observed variables;  $\mathbf{A}$ , the mixing matrix;  $\mathbf{s}$ , the vector of original sources;  $\mathbf{y}$ , the vector of the independent components; and  $\mathbf{W}$ , the demixing matrix, which we assume as being invertible. Since we are ignorant of both the input and output processes and also the original sources, the ICA methodology makes several assumptions: a) both the original sources and the components  $\mathbf{y}$  are non-Gaussian and mutually independent; b) the number of observed mixtures is equal to the number of original sources, so the unknown mixing matrix is square; c) if the independent components are equal to the original sources, the mixing matrix  $\mathbf{A}$  will be the inverse of the demixing matrix  $\mathbf{W}$ :

$$\mathbf{A} = \mathbf{W}^{-1}. \quad (3)$$

Under these assumptions we can estimate both  $\mathbf{W}$  and  $\mathbf{y}$  from  $\mathbf{x}$  by looking for some components as statistically independent as possible. Thus, the objective of ICA is to find a demixing linear mapping  $\mathbf{W}$  in which the components  $\mathbf{y}$  would be as statistically independent as possible.

In relevant literature we can find mainly three estimation criteria for ICA: a) the maximization of non-Gaussianity, b) the maximum likelihood estimation, and c) the minimization of mutual information. As it is expressed in [23], under some conditions, the three approaches are essentially equivalent or at least closely related.

The former three criteria allow for different methods of computing the ICs, which resemble one another in the sense that the optimization step is done by means of an iterative algorithm. The two main methods are: the adaptive algorithms based on gradient methods, and the fixed-point iteration scheme algorithm, known as fast fixed-point or Fast-ICA algorithm.

## 2.2 PCA, FA, ICA and Finance

In reference to PCA and FA, [21] state that ICA is capable of finding the underlying factor when these techniques fail; furthermore, [39] declare that ICA might reveal some features that otherwise would remain hidden. In addition, PCA and FA present a limitation that ICA overcomes. It is often believed that PCA and FA generate independent components; however, this is only true if the data are multivariate normally distributed, since uncorrelated components are also independent for Gaussian data.

The real world data and specially the financial time series usually are non-Gaussian. ICA will search statistically independent components for non-Gaussian data. Moreover, independence represents a stronger property than uncorrelatedness, since the former implies the latter but not vice versa. Therefore, uncorrelatedness is not enough to separate the underlying components. From a different perspective, PCA and FA techniques use only the covariance matrix to obtain linear decorrelated components, i.e., they minimize second-order statistics.

ICA uses statistics that are not considered in the covariance matrix, i.e., they additionally minimize higher-order statistics containing information not included in the covariance matrix. Consequently, another problem related to the use of PCA and FA on financial time series is the fact that, in finance, probability distributions have fat tails, and therefore the outliers can distort the estimation of the parameters in both cases.

Conversely, ICA presents a special problem absent in both PCA and FA: the estimated independent components (ICs) are not explicitly ranked as in the other methods, where the factors are automatically ranked by their eigenvalues. Additionally, therefore we have to apply an algorithm able to order the ICs according to some criteria.

In the case of financial series, on the other hand, it is reasonable to assume that there is a set of independent factors that underlie the observed time series, which might be related to political, meteorological, technical, fundamental, macroeconomic, market, national or international aspects, and that ICA might be an appropriate

model to extract them. Consequently, ICA is very suitable for use on financial time series for the following reasons: first, ICA deals with the problem of blind source separation or dealing with parallel time series, like those obtained from financial variables; secondly, ICA works with non-Gaussian random variables, which are the ones most commonly found in financial data; thirdly, from statistical and financial standpoints, ICA produces more reliable underlying components or factors, since they are statistically independent and not only uncorrelated. This fact contributes directly to the aim of extracting systematic risk factors affecting the returns on equities in a multifactor asset-pricing model like the Arbitrage Pricing Theory.

### 3 Empirical Study

#### 3.1 The Data

We used four different databases formed as follows: First, for the sake of comparison with previous research [28], we ran our study over two databases consisting of 291 quotations, formed on the basis of weekly closing prices in log-returns from 20 stocks of the Mexican Stock Exchange over the period running from July 3, 2000 to January 27, 2006.<sup>2</sup> One of these two databases is stated in returns (DBWR) and the other, in excesses of the free-risk interest rate (DBWE).<sup>3</sup>

Besides, we also used two other daily databases, one expressed in returns (DBDR) and another in excesses (DBDE). The period of the daily databases, consisting of 1410 observations from 22 stocks, extended from July 3, 2000 to January, 27, 2006.<sup>4</sup>

<sup>2</sup> The criteria utilized to choose the sample of stocks for these studies have been their inclusion in the main index of the Mexican Stock Exchange (IPC) and a survival bias during the analyzed period. The period considered was defined by the available information, the terms of the IPC index's samples and the explanatory character of this study in the pre-crisis period. More recent periods will be used in future researches where we will analyze the prediction potential of this technique during other periods of time (crisis / post-crisis).

<sup>3</sup> In consistence with our previous research [28], the riskless interest rate is assumed to be equal to the government securities' daily funding interest rate published by the Bank of Mexico.

<sup>4</sup> In the same sense, as stated in our previous research [28]: "The number of assets and the periods considered were defined by the

The returns were calculated using the logarithmic returns of the stocks' closing prices, in accordance with the following expression:

$$r = \ln(p_{it}) - \ln(p_{it-1}). \quad (4)$$

Although ICA does not require time series being stationary, by using the continuous logarithmic returns analysis to compute the returns on equities as expressed in expression 4, we already are considering that the prices time series are not stationary and that a difference has been done in order to make those series stationary in mean. In addition, as the returns are differential values, the underlying mean and trend are discarded, and thus the ICA algorithm is able to capture the interactions between the different stocks at a given moment.

On the other hand, the ICA as a methodology does not require that each time series intrinsically be stationary. What ICA assumes is that the overall set of time series preserve the same kind of interactions between times series, that is, the statistics of the observations might change, but the interaction between them captured by the matrix **W** does not change.

Finally, it is a fact that by averaging over longer time intervals, such as increasing the time period from daily to weekly to monthly, gives a time series that increasingly has a lower discrepancy (see [11]); however, the discrepancies at the high values of the returns in the QQ plots with respect to a Gaussian at the level of one month, are compatible with the assumptions about non-Gaussianity needed for the ICA algorithm.

available information in accordance with a survival bias criterion. Unfortunately, since there are many gaps in the observations of several stocks in the Mexican market, it is very difficult to build a dataset of quotations which contains both a long number of observations and a large number of stocks. In our case, the 20 and 22 stocks considered represents the maximum number of shares from which we could obtain a good enough number of observations of all of them, that allowed us to build complete and homogeneous datasets for both periodicities (without missing values). This fact constitutes a very important aspect for the correct application of the extraction technique presented. In addition, we decided to use two differently structured databases in order to test the case of weekly and daily returns as well as a larger and a smaller number of observations, according to the different studies found in literature."

**Table 1.** Descriptive statistics and Jarque-Bera Test. Database of weekly returns

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
ALFAA	0.0036	0.0041	0.0619	-0.6609	7.4108	257.0801	0.0000
ARA_01	0.0049	0.0061	0.0406	-0.1335	3.5483	4.5102	0.1049
BIMBOA	0.0032	0.0019	0.0422	0.0777	4.7718	38.3563	0.0000
CIEB	-0.0019	0.0004	0.0505	-0.7843	6.2150	155.1639	0.0000
COMERUBC	0.0023	0.0010	0.0454	0.1356	4.4699	27.0904	0.0000
CONTAL_01	0.0020	0.0000	0.0438	0.0716	4.6692	34.0319	0.0000
ELEKTRA_01	0.0027	0.0033	0.0569	-0.2465	4.3674	25.6200	0.0000
FEMSAUBD	0.0024	0.0017	0.0424	-0.2520	4.7448	39.9911	0.0000
GCARSOA1	0.0034	0.0062	0.0445	-0.3802	4.3096	27.8059	0.0000
GEOB	0.0082	0.0128	0.0629	-0.2622	5.1221	57.9405	0.0000
GFINBURO	0.0025	0.0031	0.0426	-0.3496	5.3609	73.5098	0.0000
GFNORTEO	0.0069	0.0077	0.0436	0.2487	4.5283	31.3195	0.0000
GMODELOC	0.0019	0.0017	0.0321	0.3192	5.2380	65.6702	0.0000
PE_OLES_01	0.0047	0.0000	0.0674	0.3414	4.3948	29.2415	0.0000
SORIANAB	0.0007	0.0000	0.0438	-0.0533	4.7728	38.2445	0.0000
TELECOA1	0.0013	0.0025	0.0444	-0.1219	3.7457	7.4627	0.0240
TELMEXL	0.0012	0.0000	0.0334	-0.5724	7.7828	293.2540	0.0000
TLEVICPO	0.0009	0.0020	0.0475	-0.3993	5.7427	98.9405	0.0000
TVAZTCPO	-0.0003	0.0000	0.0528	-0.3567	4.4700	32.3714	0.0000
WALMEXV	0.0033	0.0030	0.0398	-0.0261	4.5949	30.8752	0.0000

### 3.2 Methodology and Results

#### 3.2.1 Tests for Univariate and Multivariate Normality

It is known [21] that PCA (implicitly) and FA (explicitly) require a normally distributed multivariate sample in order to produce completely reliable results, i.e., they will only produce uncorrelated and independent components if the sample data have no higher order statistics beyond the variance.

Thus, if the samples do not fulfill these conditions, we will be prompted to use a more suitable technique such as ICA to uncover the underlying sources in a non-Gaussian sample. Therefore, we first tested the univariate normality

(UVN) of each individual series, since ICA requires that not more than one of the observed signals (the returns on equities) be non-Gaussian.

Tables 1 to 4 present the descriptive statistics up to the fourth moment of the four databases used in this study. We can observe that the skewness and the kurtosis of practically all the stocks differs from those of the Gaussian distribution.

We also carried out the Jarque-Bera test for UVN on the four databases, rejecting the null hypothesis of normality at 5% of probability for all the stocks in the daily databases, but not rejecting it for only one stock in the weekly databases that was normally distributed. The last two columns of the Tables 1 to 4 present the results of the Jarque-Bera test.

We used two classical alternatives for assessing the multivariate normality (MVN) tests:

**Table 2.** Descriptive statistics and Jarque-Bera Test. Database of weekly excesses

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
ALFAA	0.0019	0.0030	0.0620	-0.6709	7.3742	253.8279	0.0000
ARA_01	0.0032	0.0045	0.0406	-0.1423	3.5319	4.4115	0.1102
BIMBOA	0.0015	0.0002	0.0422	0.0699	4.7836	38.8079	0.0000
CIEB	-0.0036	-0.0010	0.0506	-0.7874	6.1942	153.7829	0.0000
COMERUBC	0.0006	-0.0005	0.0455	0.1275	4.4335	25.7027	0.0000
CONTAL_01	0.0004	-0.0018	0.0438	0.0597	4.6472	33.0725	0.0000
ELEKTRA_01	0.0010	0.0017	0.0569	-0.2500	4.3482	25.0695	0.0000
FEMSAUBD	0.0007	0.0003	0.0424	-0.2723	4.7356	40.1191	0.0000
GCARSOA1	0.0017	0.0052	0.0446	-0.4009	4.3393	29.5442	0.0000
GEOB	0.0065	0.0103	0.0630	-0.2847	5.1160	58.2218	0.0000
GFINBURO	0.0008	0.0015	0.0426	-0.3555	5.3354	72.2614	0.0000
GFNORTEO	0.0052	0.0062	0.0437	0.2379	4.4759	29.1582	0.0000
GMODELOC	0.0002	0.0001	0.0322	0.2873	5.2272	64.1473	0.0000
PE_OLES_01	0.0030	-0.0017	0.0675	0.3316	4.3801	28.4267	0.0000
SORIANAB	-0.0009	-0.0010	0.0439	-0.0721	4.7767	38.5244	0.0000
TELECOA1	-0.0004	0.0006	0.0445	-0.1458	3.7462	7.7812	0.0204
TELMEXL	-0.0005	-0.0015	0.0335	-0.6063	7.8238	299.9606	0.0000
TLEVICPO	-0.0008	0.0007	0.0476	-0.4135	5.7603	100.6749	0.0000
TVAZTCPO	-0.0020	-0.0009	0.0528	-0.3650	4.4637	32.4391	0.0000
WALMEXV	0.0016	0.0016	0.0399	-0.0627	4.5845	30.6314	0.0000

the Mardia [33] and the Henze-Zirkler [18] MVN tests. Mardia's test is based on the multivariate skewness and kurtosis of the sample. Henze-Zirkler's (H-Z) test considers a measure of the distance between the characteristic function of the MVN and the empirical one, where the computed statistic will be lognormally distributed, if the data is multivariate normal. Both techniques have shown very good performance in measuring the MVN against other classic and newer alternatives, as [34] remark in their study.

We performed two tests following the accepted criterion of applying more than one MVN test when assessing this property of a sample.<sup>5</sup> Our results with both tests reject the null hypothesis of MVN at

<sup>5</sup> We performed both MVN tests using the Matlab scripts developed by [41, 42].

5% of probability for all the databases. Tables 5 and 6 present the results of Mardia's and H-Z's tests, respectively.

We extended this analysis by making an experiment concerning the horizon of Mardia's test, i.e., we ran the test using different numbers of observations so as to check the multivariate normality in different scenarios. The results showed that from 101 observations on, inclusive, the sample is non-Gaussian according to the three statistics.

On the basis of the foregoing results<sup>6</sup>, we cannot accept as completely reliable the outcomes of techniques assuming the multivariate normality of data such as PCA and FA; thus, we are led to the

<sup>6</sup> The fact that the results of kurtosis are positive and large, revealing the presence of outliers, will have implications on the election of the non-linearity in the ICA estimation.

**Table 3.** Descriptive statistics and Jarque-Bera Test. Database of daily returns

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
ALFAA	0.0007	0.0000	0.0246	-0.1153	6.3963	680.8083	0.0000
ARA_01	0.0010	0.0000	0.0189	-0.0442	5.9361	506.9414	0.0000
BIMBOA	0.0007	0.0000	0.0187	0.3740	7.6206	1287.2010	0.0000
CIEB	-0.0004	0.0000	0.0213	-0.6673	9.9616	2951.9139	0.0000
COMERUBC	0.0005	0.0000	0.0204	0.4306	6.4539	744.4508	0.0000
CONTAL_01	0.0004	0.0000	0.0211	-0.1938	6.8047	859.2542	0.0000
ELEKTRA_01	0.0005	0.0002	0.0245	-0.1246	6.4904	719.3973	0.0000
FEMSAUBD	0.0005	0.0000	0.0175	-0.2518	7.1901	1046.3697	0.0000
GCARSOA1	0.0007	0.0000	0.0192	-0.2304	6.1817	607.2330	0.0000
GEOB	0.0017	0.0000	0.0245	-0.1054	10.2044	3051.9052	0.0000
GFINBURO	0.0005	0.0000	0.0194	0.2199	5.0447	256.9903	0.0000
GFNORTEO	0.0014	0.0000	0.0205	0.2748	6.7824	858.2517	0.0000
GMODELOC	0.0004	0.0000	0.0158	0.1737	5.6468	418.6632	0.0000
PE_OLES_01	0.0010	0.0000	0.0295	-0.3729	10.1686	3051.7488	0.0000
SORIANAB	0.0002	0.0000	0.0186	-0.0839	4.6112	154.1588	0.0000
TELECOA1	0.0003	0.0006	0.0195	-0.1156	4.7901	191.3930	0.0000
TELMEXL	0.0002	0.0000	0.0156	-0.1018	6.0378	544.6098	0.0000
TLEVICPO	0.0002	0.0006	0.0220	-0.1052	6.6617	790.3090	0.0000
TVAZTCPO	-0.0001	0.0000	0.0244	-0.5064	8.0397	1552.4342	0.0000
WALMEXV	0.0007	0.0006	0.0187	0.1244	5.9440	512.8407	0.0000
CEMEXCP	0.0008	0.0000	0.0162	0.1342	4.2068	89.7969	0.0000
KIMBERA	0.0002	0.0000	0.0151	-0.5530	9.0290	2207.3787	0.0000

application of more suitable techniques like ICA. In fact, this part of our investigation represents an important, but in most cases ignored, aspect in empiric studies that uses classic multivariate techniques to extract the pervasive factors; since in many cases the MVN is assumed but not tested, the results and conclusions may be flawed.

In addition, the assumption done in the ICA models, is that the third and fourth moments differ significantly from the values of a Gaussian distribution.

In addition, the tests of normality are based on checking this assumption. In particular the non-

linearities used for the implementation of the experiments in this paper, guaranteed the presence of high order interactions from the Taylor expansion, and therefore the presence of moments of all orders.

### 3.2.2 Estimation of the ICA Model

In order to estimate the ICA model in expression (2), we used the ICASSO methodology [20], which is based on the FastICA algorithm [22]<sup>7</sup>. According to the foregoing authors, the FastICA algorithm is based on a fixed-point iteration scheme for finding the local extrema of the objective functions. The

<sup>7</sup> We used the Matlab package developed by [19] to estimate the ICA model using the ICASSO methodology. At the same time the ICASSO

software uses the FastICA Matlab package by [13] to estimate the FastICA algorithm.

**Table 4.** Descriptive statistics and Jarque-Bera Test. Database of daily excesses

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability
ALFAA	0.0005	-0.0001	0.0246	-0.1215	6.3955	680.8189	0.0000
ARA_01	0.0008	-0.0002	0.0189	-0.0495	5.9402	508.4618	0.0000
BIMBOA	0.0004	-0.0002	0.0187	0.3744	7.6211	1287.5568	0.0000
CIEB	-0.0006	-0.0002	0.0213	-0.6697	9.9707	2960.0790	0.0000
COMERUBC	0.0003	-0.0002	0.0204	0.4273	6.4467	740.8504	0.0000
CONTAL_01	0.0002	-0.0002	0.0211	-0.1962	6.7999	857.3613	0.0000
ELEKTRA_01	0.0003	0.0000	0.0245	-0.1266	6.4854	717.4653	0.0000
FEMSAUBD	0.0002	-0.0002	0.0175	-0.2567	7.2068	1055.2038	0.0000
GCARSOA1	0.0005	-0.0001	0.0192	-0.2365	6.1774	606.2876	0.0000
GEOB	0.0015	-0.0001	0.0245	-0.1144	10.1975	3046.6028	0.0000
GFINBURO	0.0003	-0.0002	0.0193	0.2208	5.0571	260.0685	0.0000
GFNORTEO	0.0012	-0.0001	0.0205	0.2716	6.7766	855.2821	0.0000
GMODELOC	0.0001	-0.0002	0.0158	0.1670	5.6406	416.2018	0.0000
PE_OLES_01	0.0008	-0.0002	0.0295	-0.3695	10.1326	3020.9541	0.0000
SORIANAB	-0.0001	-0.0002	0.0186	-0.0883	4.6225	156.4975	0.0000
TELECOA1	0.0000	0.0005	0.0195	-0.1242	4.7890	191.6613	0.0000
TELMEXL	0.0000	-0.0002	0.0156	-0.1130	6.0560	551.6562	0.0000
TLEVICPO	-0.0001	0.0004	0.0220	-0.1122	6.6667	792.8200	0.0000
TVAZTCPO	-0.0003	-0.0002	0.0244	-0.5083	8.0248	1544.0783	0.0000
WALMEXV	0.0004	0.0004	0.0187	0.1142	5.9465	513.1155	0.0000
CEMEXCP	0.0006	-0.0002	0.0161	0.1316	4.2152	90.8231	0.0000
KIMBERA	0.0000	-0.0002	0.0151	-0.5621	9.0350	2213.9756	0.0000

basic iteration for the vector  $\mathbf{w}$  for each IC obtained by this method is:

$$\mathbf{w} \leftarrow E\{\mathbf{z}g(\mathbf{w}^T\mathbf{z})\} - E\{g'(\mathbf{w}^T\mathbf{z})\}\mathbf{w}. \quad (5)$$

where the nonlinearity  $g$  can be almost any smooth function such as:

$$g_1(y) = \tanh(a_1 y). \quad (6)$$

$$g_2(y) = y \exp(-y^2/2). \quad (7)$$

$$g_3(y) = y^3. \quad (8)$$

<sup>8</sup> According to [21], nonlinearity  $\tanh(a_1 y)$  is optimal for super-Gaussian fat-tail distributions;  $y^3$  performs better for sub-Gaussian thin-

and  $g'$  is the derivative of  $g(\cdot)$ .<sup>8</sup>

The final vector gives one of the ICs as a linear combination in  $\mathbf{y} = \mathbf{w}^T\mathbf{z}$ . The specific resulting algorithm depends both on the estimation principle used and the approach selected to estimate several numbers of ICs, i.e., the nonlinearity and the decorrelation method chosen. In [21], the authors state that by setting the options, nonlinearity  $\tanh$  (hyperbolic tangent) and symmetric approach, one can obtain a good estimation of the ICA model; this would be

tail ones; and  $y \exp(+y^2/2)$  is recommended for highly super-Gaussian distributions or when robustness is very important.



**Table 5.** Mardia Test for Multivariate Normality

	DBWR	DBWE	DBDR	DBDE
Multivariate Skewness (Ms)	3305.50	3297.10	6659.40	6666.30
p-value	0.00	0.00	0.00	0.00
Multivariate Skewness corrected (Msc)	3342.80	3334.40	6674.80	6681.70
p-value	0.00	0.00	0.00	0.00
Multivariate Kurtosis (Mk)	37.83	37.71	141.05	141.16
p-value	0.00	0.00	0.00	0.00

**Notes:**

DBWR = Database of weekly returns. DBWE = Database of weekly excesses. DBDR= Database of daily returns. DBDE= Database of daily excesses.  $H_0$  = Multivariate Normality. p-value lower than 0.05 = Rejection of the  $H_0$ .

**Table 6.** Henze-Zirkler Test for Multivariate Normality

	DBWR	DBWE	DBDR	DBDE
Henze-Zirkler's Statistic	1.05	1.05	1.22	1.22
p-value	0.00	0.00	0.00	0.00

**Notes:**

DBWR = Database of weekly returns. DBWE = Database of weekly excesses. DBDR= Database of daily returns. DBDE= Database of daily excesses.  $H_0$  = Multivariate Normality. p-value lower than 0.05 = Rejection of the  $H_0$ .

equivalent to performing the three estimation approaches at the same time.

In addition, the positive kurtosis obtained in the multivariate normality tests leads us to use the hyperbolic tangent function.

Furthermore, as reported in [14], the best trade-off for estimating the ICA model, from statistical performance and computational load perspectives, is represented by the FastICA algorithm with symmetric orthogonalization and *tanh* nonlinearity estimation. In our study we followed these specifications.

The election of the ideal number of ICs to estimate still represents an unsolved problem.

Although in ICA literature we can find diverse criteria to determine this number, in most cases it is actually chosen by trial and error without any theoretical basis. One alternative is to reduce the number of dimensions in the whitening pre-

processing stage, considering some criteria from among those used in PCA or FA, and to estimate the same number of ICs. For the sake of comparison with our previous study, we use the same test window, which ranges from two to nine components.<sup>9</sup>

As stated by [20], one problem that the ICA estimation presents is that the reliability of the estimated ICs is not known since the results are stochastic, i.e., they might be dissimilar in different runs of the algorithm.

Thus, the results of a single run of the FastICA algorithm could not be completely trusted and an additional analysis of the reliability of the estimation should be performed. In this context, reliability has two aspects the algorithmic and the statistical. According to the former authors, ICASSO methodology represents an alternative for dealing with this problem, since it ensures the

<sup>9</sup> The criteria adopted were the same used in our previous research [28]: "the arithmetic mean of the eigenvalues, the percentage of explained variance, the exclusion of the components or factors explaining a small amount of variance, the scree plot, the unretained eigenvalue contrast (Q statistic), the likelihood ratio contrast, Akaike's

information criterion (AIC), the Bayesian information criterion (BIC), and the maximum number of components feasible to estimate in each technique."

algorithmic and statistical stability and reliability of the estimated components by running the FastICA algorithm many times, using different initial conditions and/or a differently bootstrapped data set.

Following [20], ICASSO first runs the FastICA algorithm  $M$  times on data set  $X = [x_1, x_2, \dots, x_N]$ , composed of  $N$  samples of  $k$  vectors; then, ICASSO forms clusters with the ICs produced in each run according to their similarity. Mutual similarities between estimates are computed, using the absolute value of their linear correlation coefficient as the measure of similarity:

$$\sigma_{ij} = |r_{ij}|. \quad (9)$$

These elements form the similarity matrix, which can be obtained by:

$$\mathbf{R} = \widehat{\mathbf{W}} \mathbf{\Sigma} \widehat{\mathbf{W}}^T, \quad (10)$$

where,  $\mathbf{\Sigma}$  is the covariance matrix of dataset  $\mathbf{x}$ , and  $\widehat{\mathbf{W}}$  is the estimates of demixing matrices  $\widehat{\mathbf{W}}_i$  from each run  $i = 1, 2, \dots, M$  gathered in a single matrix:

$$\widehat{\mathbf{W}} = [\widehat{\mathbf{W}}_1^T \widehat{\mathbf{W}}_2^T \dots \widehat{\mathbf{W}}_M^T]. \quad (11)$$

According to [19], reliable estimates of ICs correspond to tight clusters, since they agglomerate estimates generated by many runs of the algorithm which are similar, even when the initial values and datasets for the estimation have been changed. Conversely, estimates which do not belong to any cluster are considered unreliable estimates. The centrotpe of each cluster is considered a more reliable estimate than that generated by any single run.

Besides the previously declared parameters for FastICA, there are some additional parameters to set when using ICASSO, such as the resampling mode, number of resampling cycles ( $M$ ) and number of clusters ( $L$ ). In order to ensure both statistical and algorithmic reliability, in our study we used both resampling modes, i.e., each time the dataset was bootstrapped and the initial conditions of the algorithm were randomized. We used the default number of resampling cycles fixed by the software, i.e., 30, and we set the number of clusters according to the number of ICs ( $m$ ) estimated in each experiment in order to obtain squared mixing ( $\mathbf{A}$ ) and demixing ( $\mathbf{W}$ ) matrices.

The demixing matrix ( $\mathbf{W}$ ) computed by ICASSO corresponds to the centrotypes of each cluster as well, representing a more reliable estimate than that produced by a single run of FastICA; however, they are not strictly orthogonalized. In the context of our research where we need to obtain orthogonalized ICs, we will have to make an orthogonalization procedure in a later step.

Consequently, we first took the demixing matrix ( $\mathbf{W}$ ) produced by ICASSO, then we computed the mixing matrix:

$$\mathbf{A} = \mathbf{W}^{-1}, \quad (12)$$

and the matrix of independent components or sources:

$$\mathbf{S} = \mathbf{W}\mathbf{X}. \quad (13)$$

### 3.2.3 Ranking and Orthogonalization of the Independent Components

The ICA model does a decomposition by means of a criterion related to statistical independence, which does not allow to order in a natural way the components and thus the residual. The criterion presented in this section is one criterion that has sense in the application at hand. In contrast with the case of linear regression or PCA, where the driving noise is easy to identify, because it is a residual obtained after the components of maximum variance are determined, in the case of ICA such an interpretation will not be natural. Because of this, in the literature about ICA it is not clearly specified the difference between the components and the residual, and therefore the results are usually presented as a complete projection in the space statistically independent components.

Then, next we ordered the independent components in terms of their explained variability by means of the criterion proposed by [12]. This criterion ranks the ICs according to the amount of variance of the stocks that explains each one of them, thus we obtain a ranked matrix of independent components ( $\mathbf{S}'$ ), as well as sorted mixing ( $\mathbf{A}'$ ) and demixing matrices ( $\mathbf{W}'$ ).

Finally, we orthogonalized the matrix of ICs by means of the following process of transformation:

$$\mathbf{V} = 2 * ((\mathbf{S}'^T * \mathbf{S}')^{-1})^{1/2}, \quad (14)$$

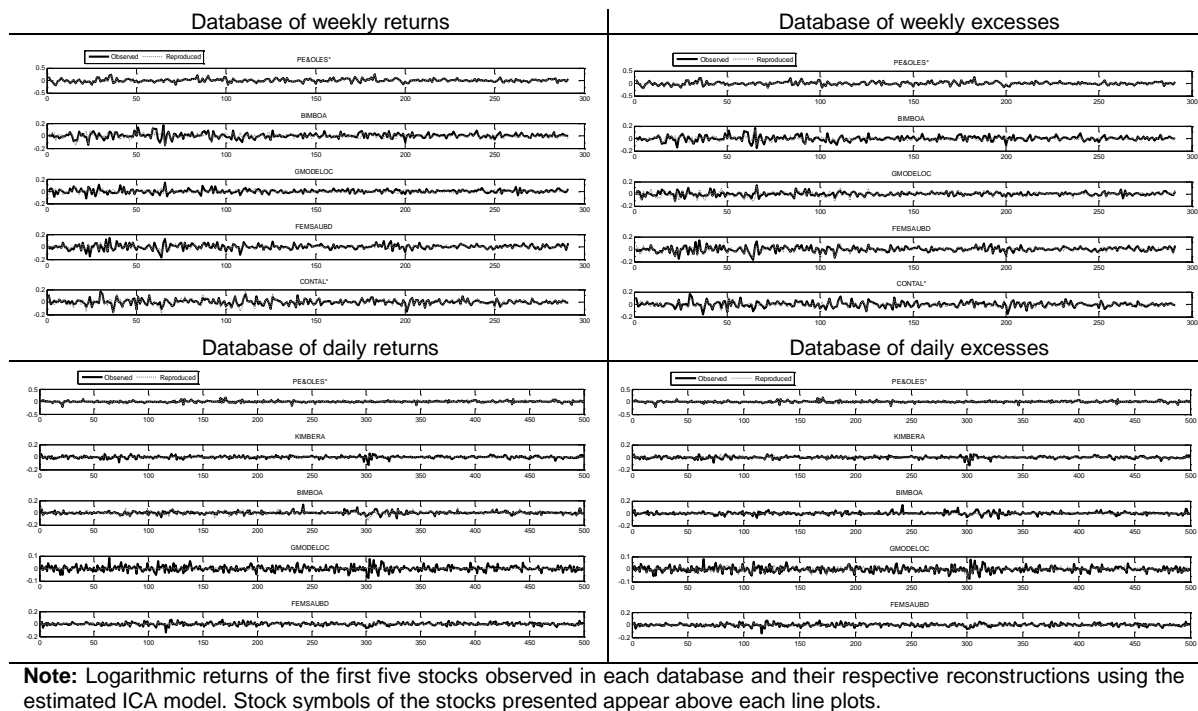


Fig. 1. Line plots of the observed and reproduced stocks

$$\mathbf{S}^0 = \mathbf{V} * \mathbf{S}^r, \quad (15)$$

where  $\mathbf{V}$  is a transformation matrix to decorrelate the matrix of sorted independent components, and  $\mathbf{S}^0$  represents the matrix of orthogonalized ICs.

### 3.2.4 Extraction of Underlying Systematic Risk Factors Via ICA

In each one of the four databases, we computed eight multifactor models in order to extract a window from two to nine independent components. Then, we proceeded to reconstruct the original variables according to the generation process of expression (1), including the inverse of the transformation matrix  $\mathbf{V}$  in order to orthogonalize the mixing matrix  $\mathbf{A}$  as well:

$$\mathbf{X} = \mathbf{S}^0 (\mathbf{V}^{-1} * \mathbf{A}^r). \quad (16)$$

The reproduced values were very similar to the observed series for greater part of the equities in

all the datasets, which indicates that the generative multifactor model performed by ICA was effective. However, stocks such as GMODELO, CEMEX, SORIANA and GCARSO were not very well reconstructed, especially in the cases of daily returns and excesses, due to the high volatility they presented during the studied period. To save space, we only present the line plots for the first five stocks appearing in the returns and excesses observed and reproduced from each database.

Figures 1 to 4 present the results of the case when we extracted nine underlying factors; the reconstruction performance is evident.<sup>10</sup> An interesting fact of the ICA algorithm is that it captures the global interaction between stocks, independently of the non-stationarity of the joint behavior. That is, the required assumption in the model is that there are independent sources that are mixed by a matrix  $\mathbf{W}$ .

If the matrix does not change, the ICA algorithm will give an estimation, and therefore, given that

<sup>10</sup> As in our previous paper [28], the rest of the estimations when we extract 2, 3, 4, 5, 6, 7 and 8 components showed similar behavior. The observed results are typical.

the matrix does not change, it will impute the components of volatility to some of the non-observable factors.

### 3.2.5 Independence Test

In order to test the independence of the computed ICs, we ran the Hilbert-Schmidt Independence Criterion (HSIC) test [15]<sup>11</sup>, which tests whether random variables  $X$  and  $Y$  are independent based on a sample of observed pairs  $(x_i, y_i)$ . The results of our independence tests confirmed the statistical independence, between each pair of components estimated from the weekly and daily databases.

### 3.2.6 Econometric Contrast

We carried out an econometric contrast under a statistical approach to the Arbitrage Pricing Theory (APT) using the underlying systematic risk factors extracted via ICA. The APT's pricing equation is expressed as follows:

$$E(R_i) = \lambda_0 + \lambda_1 \cdot \beta_{1i} + \lambda_2 \cdot \beta_2 + \dots + \lambda_k \cdot \beta_{ki} \quad (17)$$

In the same outline that in [28],  $\lambda_0$  represents the riskless interest rate,  $\lambda_k$  the risk premium for each kind of systematic risk factor, and  $\beta_k$  the exposures to each type of systematic risk. We tested the former expression by way of an average cross-section methodology estimating the coefficients by ordinary least squares (OLS) in the following regression model:

$$\bar{R}_i = \lambda_0 + \lambda_1 \cdot \beta_1 + \lambda_2 \cdot \beta_2 + \dots + \lambda_k \cdot \beta_{ki} + \bar{\epsilon}_i \quad (18)$$

We used again the two-stage methodology for the econometric contrast of the APT used in our aforementioned study [28], which is explained as follows: In the first stage, we estimated the betas to be used in expression 18 from the scores of the extracted factor. In the second stage, we estimated the lambdas. In the first stage we estimated the betas by regressing the factor scores obtained by ICA as a cross-section on the returns and excesses. In order to improve the efficiency of the parameter estimates and to eliminate autocorrelation in the error terms of the regressions, we used weighted least squares

(WLS) to estimate the entire system of equations at the same time.

The results of the regressions in the four databases were very good, producing, in almost all cases, statistically significant parameters, high values of the  $R^2$  coefficients and results in the Durbin-Watson test of autocorrelation, which lead us to the non-rejection of the null hypothesis of no-autocorrelation. In the second stage we estimated the lambdas or risk premia in expression 17 by regressing the betas obtained in the first stage as a cross-section on the average returns and excesses, using ordinary least squares (OLS).

In order to avoid the econometric problems of heteroskedasticity and autocorrelation in the residuals of the model estimated through OLS, we corrected it by means of the Newey-West heteroskedasticity and autocorrelation consistent covariance estimates (HEC). Additionally, we verified the normality in the residuals by carrying out the Jarque-Bera test of normality.

In order to accept the APT pricing model, we require the statistical significance of at least one parameter lambda different from  $\lambda_0$ , and the equality of the independent term to its theoretic value, i.e., the average returns, in the models expressed in returns:

$$\lambda_0 = \bar{R}_0, \quad (19)$$

and zero, in the models expressed in excesses of the riskless interest rate:

$$\lambda_0 = 0. \quad (20)$$

We used Wald's test to confirm these equalities.

In Table 7, we present a summary of the results of the econometric contrast for the four databases. In general, the results of the explanation power, the adjusted R-squared ( $R^2$ ), the statistical significance of the multivariate test ( $F$ ), and the Jarque-Bera normality test of the residuals are very good in almost all the contrasted models. The univariate tests for the individual statistical significance of the parameters (statistic  $t$ ) priced from one to five factors exclusive of  $\lambda_0$  in the weekly and daily databases, thus giving evidence in favor of the APT in 27 models.

<sup>11</sup> We performed HSIC test using the Matlab script developed by [16].

Table 7. Summary of the Econometric Contrast

	$\lambda_0$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$\lambda_8$	$\lambda_9$	$R^2$	$\lambda_{sig} / \lambda_{tot}$	$t$	$W$	$J$
													ALD		-B
<b>Database of weekly returns.</b>															
Model with 2 betas	•	•	•								5.7	0.0	•	○	○
Model with 3 betas	0.0	•	•	0.0							8%	0%	•	•	○
Model with 4 betas	0530	•	-	1665	•						46.	33%	•	•	○
Model with 5 betas	0.0	•	-	-	•						46.	50.	•	•	○
Model with 6 betas	0546	•	0.01492	0.01219	•	•					58%	00%	•	•	○
Model with 7 betas	0.0	•	-	•	•	•					47.	20.	•	•	○
Model with 8 betas	0507	•	0.01770	-	•	•	•				28%	00%	•	•	○
Model with 9 betas	0.0	•	-	•	•	•	•	•			44.	16.	•	•	○
Model with 10 betas	0546	•	0.01899	•	•	•	•	•	•		21%	67%	•	•	○
Model with 11 betas	0.0	•	0.0	•	•	•	•	•	•	•	38.	14.	•	•	•
Model with 12 betas	0505	•	2035	•	•	•	•	•	•	•	45%	29%	•	•	•
Model with 13 betas	0.0	•	0.0	-	•	•	•	•	•	•	49.	25.	•	•	○
Model with 14 betas	0557	•	1043	0.01765	•	•	•	•	•	•	69%	00%	•	•	○
Model with 15 betas	0.0	•	•	•	•	-	•	•	•	•	34.	11.	•	•	○
Model with 16 betas	0557	•	•	•	0.01158	•	•	•	•	•	51%	11%	•	•	○
<b>Database of weekly excesses.</b>															
Model with 2 betas	•	•	•								17.	0.0	•	○	○
Model with 3 betas	0.0	•	•	0.0							81%	0%	•	•	○
Model with 4 betas	0376	•	-	1662	•						37.	33.	•	•	○
Model with 5 betas	0.0	•	•	0.0	•						21%	33%	•	•	○
Model with 6 betas	0341	•	0.01774	0890	•	•					45.	50.	•	•	○
Model with 7 betas	•	•	•	•	•	•					25%	00%	•	•	○
Model with 8 betas	•	•	•	•	•	•	•				-	0.0	•	•	○
Model with 9 betas	0.0	•	•	•	•	•	•	•			29.79%	0%	•	•	○
Model with 10 betas	0249	•	•	•	•	•	0.0	•	•		39.	16.	•	•	○
Model with 11 betas	•	•	•	•	•	•	1716	•	•	•	81%	67%	•	•	○
Model with 12 betas	•	•	•	•	•	•	0.0	-	•	•	31.	14.	•	•	○
Model with 13 betas	•	•	•	•	•	•	1431	0.00499	•	•	63%	29%	•	•	○
Model with 14 betas	•	•	•	•	•	•	•	•	0.01046	•	9.3	12.	•	•	○
Model with 15 betas	0.0	•	-	•	•	0.0	•	0.0	-	0.0	4%	50%	•	•	○
Model with 16 betas	0450	•	0.01257	•	•	1049	•	1246	0.01057	0941	63.	55.	•	•	○
Model with 17 betas	•	•	•	•	•	•	•	•	•	•	49%	56%	•	•	○
<b>Database of daily returns.</b>															
Model with 2 betas	•	•	•								-	0.0	•	○	○
Model with 3 betas	0.0	•	•	•							2.48%	0%	•	○	○
Model with 4 betas	0055	•	0.00302	•	•						30.	33.	•	•	○
Model with 5 betas	0.0	•	0.0	-	•						49%	33%	•	•	○
Model with 6 betas	0108	•	0286	0.00262	•	•					52.	50.	•	•	○
Model with 7 betas	0.0	•	•	•	•	•					34%	00%	•	•	○
Model with 8 betas	0.0	•	•	•	•	•	•				46.	20.	•	•	○
Model with 9 betas	0105	•	0.00254	•	•	•	•	•			41%	00%	•	•	○
Model with 10 betas	•	•	•	•	•	0.0	-	•	•		40.	33.	•	•	○
Model with 11 betas	•	•	•	•	•	0.0290	0.00162	•	•	•	33%	33%	•	•	○
Model with 12 betas	•	•	•	•	0.0	•	0.0	•	•	•	40.	28.	•	•	○
Model with 13 betas	0.0	0.0	0.0	•	0.0	•	0.0	•	•	0.00	22%	57%	•	•	○
Model with 14 betas	0131	0243	0329	•	0.0	•	•	•	0.00	2665	56.	50.	•	•	○
Model with 15 betas	•	•	•	•	•	•	•	•	•	•	08%	00%	•	•	○
Model with 16 betas	•	•	0.00353	•	•	0.0	•	•	•	•	69.	33.	•	•	○
Model with 17 betas	•	•	•	•	•	0.0287	•	•	•	01	62%	33%	•	•	○
<b>Database of daily excesses.</b>															
Model with 2 betas	•	•	•								-	0.0	•	○	○
Model with 3 betas	•	•	0.0	•							1.91%	0%	•	○	○
Model with 4 betas	•	•	0318	•	•						34.	33.	•	•	○
Model with 5 betas	•	•	•	0.0	•						55%	33%	•	•	○
Model with 6 betas	•	•	•	0.0	•	•					50.	25.	•	•	○
Model with 7 betas	•	•	•	0244	•	•	•				53%	00%	•	•	○
Model with 8 betas	•	•	•	•	•	•	•	•			39.	20.	•	•	○
Model with 9 betas	•	•	0.00289	•	•	•	•	•	•		87%	00%	•	•	○
Model with 10 betas	•	•	•	•	0.0	•	•	•	•	•	36.	16.	•	•	○
Model with 11 betas	•	•	•	•	0.0	•	•	•	•	•	25%	67%	•	•	○
Model with 12 betas	•	•	0.0	•	•	•	•	•	•	•	45.	28.	•	•	○
Model with 13 betas	•	•	0.0	•	•	•	•	•	•	•	30%	57%	•	•	○
Model with 14 betas	•	-	•	•	0.0	•	0.0	•	•	•	44.	37.	•	•	○
Model with 15 betas	0.00197	•	•	•	0.0	•	0.0	•	•	•	95%	50%	•	•	○
Model with 16 betas	•	0.0	-	•	0.0	•	•	•	0.00	0.0	78.	66.	•	•	○
Model with 17 betas	0300	0.00183	0250	•	•	0.00076	•	•	2742	0109	98%	67%	•	•	○

Notes: (1) The level of statistical significance used in all the tests was 5%. (2) Empty circles mean that the required results in the different tests were fulfilled, whereas filled circles represent that those tests were not passed according to the different null hypotheses posed in each one of them. (3)  $\lambda_i$ : Estimated coefficients.  $H_0: \lambda_i = 0$ . Numeric value of the coefficient = Rejection of  $H_0$ . Parameter significant. • = Not rejection of  $H_0$ . Parameter not significant. (4)  $R^2$ : Adjusted R-squared = Explanatory capacity of the model. (5)  $\lambda_{sig} / \lambda_{tot}$ : Ratio number of significant lambdas / total number of lambdas in the model. (6)  $F$ : Global statistical significance of the model.  $H_0: \lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ . • = Rejection of  $H_0$ . Model globally significant. • = Not rejection of  $H_0$ . Model globally not significant. (7)  $W$ : Wald's test for coefficient restrictions. Databases in returns:  $H_0: \lambda_0 = \text{Average riskless interest rate}$ . Databases in excesses:  $H_0: \lambda_0 = 0$ . • = Not rejection of  $H_0$ . The independent term is equal to its theoretic value. • = Rejection of  $H_0$ . The independent term is not equal to its theoretic value. (8)  $J-B$ : Jarque Bera's test for normality of the residuals.  $H_0 = \text{Normality}$ . • = Not rejection of  $H_0$ . The residuals are normally distributed. • = Rejection of  $H_0$ . The residuals are not normally distributed.

Nevertheless, only four models fulfilled both the statistical significance of the parameters and the equality of the independent term to its theoretic value, in addition to the fulfilment of normality in the residuals.

The referred models appear marked in Table 7, where we used the same methodology of presentation and analysis of the results as in our preceding paper [28].

## 4 Conclusions

Our results showed that the data of the Mexican Stock Exchange used in the study presented univariate and multivariate non-Gaussianity, revealing that classic techniques such as PCA and FA will produce a biased estimation of the betas.

This discovery led us directly to the use of techniques more suitable for non-Gaussian series such as ICA, which, by using the ICASSO methodology, produces a more reliable and realistic estimation of the underlying generative multifactor model of returns on equities than those produced by PCA and FA, since this methodology is capable of extracting the underlying systematic risk factors from non-Gaussian financial time series, and solves the problem that the regular ICA model estimation presents.

Regarding the results of our empirical study, on one hand, the reconstruction of the observed signals, by means of a reduced number of factors with respect to the original variables with our estimated ICA model was suitable. On the other hand, our econometric contrast of the APT in the stocks and periods used in this study produced signals in favor of the APT, revealing from 1 to 5 factors priced in the statistically significant models.

Compared with the results of our previous study [28] and given the univariate and multivariate non-gaussianity of the financial time series used in both studies, we find that from a theoretical standpoint, the underlying systematic factors extracted using ICA would represent a more reliable estimation than that produced by PCA and FA. Nevertheless, from an empirical stance, in general, both the reconstruction of the observed data and the results of the econometric contrast of the APT were similar. Further research will be needed in order to

compare the performance of these extraction techniques in this context.

## Acknowledgments

The authors thank Aapo Hyvärinen from the University of Helsinki for the technical advice on some topics related to this investigation, and Cristina Urbano at Gaesco for the financial data provided.

## References

1. Ané, T. & Labidi, C. (2001). Implied volatility surfaces and market activity over time. *Journal of Economics and Finance*, Vol. 25, No. 3, pp. 259–275. DOI: 10.1007/BF02745888.
2. Back, A. & Weigend, A. (1997). A first application of independent component analysis to extracting structure from stock returns. *International Journal of Neural Systems*, Vol. 8, No. 4, pp. 473–484. DOI: 10.1142/S0129065797000458.
3. Bellini, F. & Salinelli, E. (2003). Independent component analysis and immunization: An exploratory study. *International Journal of Theoretical Applied Finance*, Vol. 6, No. 7, pp. 721–738. DOI: 10.1142/S0219024903002201.
4. Cha, S. M. & Chan, L. W. (2002). Applying Independent Component Analysis to Factor Model in Finance. K. Leung *et al.* (Eds.), *Lecture Notes in Computer Science*, Vol. 1983, pp. 538–544. DOI: 10.1007/3-540-44491-2\_78.
5. Chan, L. W. & Cha, S. M. (2001). Selection of independent factor model in finance. *Proceedings of the 3rd International Conference on ICA and Blind Signal Separation*, pp. 161–166.
6. Chen, Y., Härdle, W., & Spokoyny, V. (2007). Portfolio Value at Risk based on Independent Component Analysis. *Journal of Computational and Applied Mathematics*, Vol. 205, No. 1, pp. 594–607. DOI: 10.1016/j.cam.2006.05.016.
7. Chen, Y., Härdle, W., & Spokoyny, V. (2010). GHICA – Risk analysis with GH distributions and independent components. *Journal of Empirical Finance*, Vol. 17, No. 2, pp. 255–269. DOI: /10.1016/j.jempfin.2009.09.005.
8. Cléménçon, S. & Slim, S. (2007). On portfolio selection under extreme risk measure: The heavy-tailed ICA model. *International Journal of Theoretical and Applied Finance*, Vol. 10, No. 3, pp. 449–474. DOI: 10.1142/S0219024907004275.

9. Coli, M., Di Nisio, R., & Ippoliti, L. (2005). Exploratory analysis of financial time series using Independent Component Analysis. *Proceedings of the 27th international conference on information technology interfaces*, pp. 169–174. DOI: 10.1109/ITI.2005.1491117.
10. De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). An introduction to independent component analysis. *Journal of Chemometrics*, Vol. 14, No. 3, pp. 123–149. DOI: 10.1002/1099-128X(200005/06)14:3<123::AIDCEM589>3.0.CO;2-1.
11. Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, Vol. 38, No. 1, pp. 34–105. DOI: 10.1086/294743.
12. García-Ferrer, A., González-Prieto, E., & Peña, D. (2012). A conditional heteroskedastic independent factor model with an application to financial stock returns. *International Journal of Forecasting*, Vol. 28, No. 1, pp. 70–93. DOI:10.1016/j.ijforecast.2011.02.010.
13. Gävert, H., Hurri, J., Särelä, J., & Hyvärinen, A. (2005). *The FastICA package for Matlab*. Available at: <http://www.cis.hut.fi/projects/ica/fastica>
14. Giannakopoulos, X., Karhunen, J., & Oja, E. (1999). An experimental comparison of neural algorithms for independent component analysis and blind separation. *International Journal of Neural Systems*, Vol. 9, No. 2, pp. 99–114. DOI: 10.1142/S0129065799000101.
15. Gretton, A., Fukumizu, K., Teo, C. H., Song, L., Schölkopf, B., & Smola, A. J. (2008). Kernel Statistical Test of Independence. J. C. Platt et al. (Eds.), *Advances in Neural Information Processing Systems 20: Proceedings of the 2007 Conference*. pp. 585–592. Cambridge: MIT Press.
16. Gretton, A. (2007). *Kernel Statistical Test of Independence package for Matlab*. Available: <http://people.kyb.tuebingen.mpg.de/arthur/indep>.
17. Han, C. (2014). Measuring the dependency between securities via factor ICA models. *Journal of Applied Finance and Banking*, Vol. 4, No. 1, pp. 243–295.
18. Henze, N. & Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics -Theory Methods*, Vol. 19, No. 10, pp. 3595–3617. DOI: 10.1080/03603610929008830400.
19. Himberg, J. & Hyvärinen, A. (2005). *The ICASSO package for Matlab*. Available at: <http://research.ics.tkk.fi/ica/ICASSO/about+download.shtml>.
20. Himberg, J., Hyvärinen, A., & Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage*, Vol. 22, No. 3, pp. 1214–1222. DOI: 10.1016/j.neuroimage.2004.03.027.
21. Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. USA: Wiley-Interscience. Available at: [https://www.cs.helsinki.fi/u/ahyvarin/papers/bookfinal\\_ICA.pdf](https://www.cs.helsinki.fi/u/ahyvarin/papers/bookfinal_ICA.pdf).
22. Hyvärinen, A. & Oja, E. (1997). A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Computation*, Vol. 9, No. 7, pp. 1483–1492. DOI: 10.1162/neco.1997.9.7.1483.
23. Hyvärinen, A. & Oja, E. (2000). Independent Component Analysis: algorithms and applications. *Neural Networks*, Vol. 13, No. 4-5, pp. 411–430. DOI: 10.1016/S0893-6080(00)00026-5.
24. Korizis, H., Mitianoudis, N., & Constantinides, A. (2007). Compact representations of market securities using smooth component extraction. M. Davis et al. (Eds.), *Lectures Notes in Computer Science 4666*, pp. 738–745. DOI: 10.1007/978-3-540-74494-8\_92.
25. Kumiega, A., Neururer, T., & Van-Vliet, B. (2011). Independent Component Analysis for realized volatility: Analysis of the stock market crash of 2008. *The Quarterly Review of Economics and Finance*, Vol. 51, No. 3, pp. 292–302. DOI:10.1016/j.qref.2011.03.002.
26. Kumiega, A., Neururer, T., & Van-Vliet, B. (2012). Implied ICA: Factor extraction and multiasset derivative pricing. *The Journal of Derivatives*, Vol. 19, No. 4, pp. 39–52. DOI: 10.3905/jod.2012.19.4.039.
27. Kumiega, A., Sterijevski, G., & Vliet, B. (2014). Perspectives on hedge fund herding: A survey of analytical methods. *Wilmott*, Vol. 2014, No. 72, pp. 66–81. DOI: 10.1002/wilm.10350.
28. Ladrón de Guevara, R. & Torra, S. (2014). Estimation of the underlying structure of systematic risk with the use of principal component analysis and factor analysis. *Contaduría y Administración*, Vol. 59, No. 3, pp. 197–234. DOI: 10.1016/S0186-1042(14)71270-7.
29. Lin, T. & Chiu, S. (2013). Using independent component analysis and network DEA to improve bank performance evaluation. *Economic Modelling*, Vol. 32, pp. 608–616. DOI:10.1016/j.econmod.2013.03.003.
30. Lizieri, C., Satchell, S., & Zhang, Q. (2007). The underlying return-generating factors for REIT returns: An application of independent component analysis. *Real Estate Economics*, Vol. 35, No. 4, pp. 569–598. DOI: 10.1111/j.1540-6229.2007.00201.x.
31. Lu, C. (2010). Integrating independent component analysis-based denoising scheme with neural

- networks for stock price prediction. *Expert Systems with Applications*, Vol. 37, No. 10, pp. 7056–7054. DOI: 10.1016/j.eswa.2010.03.012.
32. **Madan, D. & Yen, J. (2008).** Asset allocation with multivariate non-Gaussian returns. In: J. Birge and V. Linetsky (Eds.), *Handbooks in Operation Research and Management Sciences*, Vol. 15, pp. 949–969. DOI: 10.1016/S0927-0507(07)15023-4.
33. **Mardia, K. (1970).** Measures of multivariate skewness and kurtosis with applications. *Biometrika*, Vol. 57, No. 3, pp. 519–530. DOI: 10.1093/biomet/57.3.519.
34. **Mecklin, C. & Mundfrom, D. (2004).** An appraisal and bibliography of tests for multivariate normality. *International Statistical Review*, Vol. 72, No. 1, pp. 123–138. DOI: 10.1111/j.1751-5823.2004.tb00228
35. **Molgedey, L. & Galic, E. (2001).** Extracting factors for interest rate scenarios. *The European Physical Journal B – Condensed Matter and Complex Systems*, Vol. 20, No. 4, pp. 517–522. DOI: 10.1007/PL00022986.
36. **Moody, J. & Yang, H. (2001).** Term Structure of Interactions of Foreign Exchange Rates. In: Y. Abu-Mostafa et al. (Eds.), *Computational Finance 1999*. Cambridge: MIT Press, pp. 247–266.
37. **Nestler, S. (2007).** Non-Gaussian asset allocation in the Federal Thrift Saving Plans. In: S. Anderson et al. (Eds.) *Proceedings of the Winter Simulation Conference*. pp. 1004–1012. DOI: 10.1109/WSC.2007.4419698.
38. **Neururer, T. & Kumiega, A. (2013).** Multifactor index variance: The case of the SPX 2000 to 2010. *The Journal of Future Markets*, Vol. 33, No. 2, pp. 158–182. DOI: 10.1002/fut.20552.
39. **Oja, E., Kiviluoto, K., & Malaroiu, S. (2000).** Independent component analysis for financial time series. *Proceedings of the IEEE Adaptive systems for signal processing, communications, and control symposium*. pp. 111–116. DOI: 10.1109/ASSPCC.2000.882456.
40. **Oja, E. (2004).** Applications of independent component analysis. *Proceedings of the International Conference on Neural Information Processing*. pp. 1044–1051. DOI: 10.1007/978-3-540-30499-9\_162.
41. **Trujillo, A. & Hernandez, R. (2003).** Mskekur: Mardia's multivariate skewness and kurtosis coefficients and its hypotheses testing.
42. **Trujillo, A., Hernández, R., Barba, K., & Cupul, L. (2007).** HZmvntest: Henze-Zirkler's Multivariate Normality Test. Available at: <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=17931>.
43. **Vermoken, M., Szafarz, A., & Pirotte, H. (2010).** Sector classification through non-Gaussian similarity. *Applied Financial Economics*, Vol. 20, No. 11, pp. 861–878. DOI: 10.1080/09603101003636238.
44. **Villavicencio, J. R., Márquez, L., & Álvarez, J. (2014).** A heuristic approach for Blind Source Separation of instant mixtures. *Computación y Sistemas*, Vol. 18, No. 4, pp. 719–730. DOI: 10.13053/CyS-18-4-1951.
45. **Wang, J., Dong, J., & Zhou, Z. (2010).** Based on Independent Component Analysis method to analyze the influence factors of close-end funds fluctuation by Shanghai stock market. J. Shaeffer (Ed.), *Proceedings of the International Conference on Management and Service Science*. pp. 1–4. DOI: 10.1109/ICMSS.2010.5576828.
46. **Wu, E. & Yu, P. (2006).** Patter recognition of the term structure using Independent Component Analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 20, No. 2, pp. 173–188. DOI: 10.1142/S0218001406004594
47. **Wu, E., Yu, P., & Li, W. (2006).** Value at Risk estimation using Independent Component Analysis-Generalized Autoregressive Conditional Heteroscedasticity (ICA-GARCH) models. *International Journal of Neural Systems*, Vol. 16, No. 5, pp. 371–382. DOI: 10.1142/S0129065706000779.
48. **Xu, Q. & Jiang, C. (2006).** Estimation for conditional higher moments risk based on Independent Component Analysis. *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics*. pp. 2358–2362. DOI: 10.1109/ICMLC.2006.258725.

Article received on 25/05/2018; accepted on 15/07/2018.  
Corresponding author is Rogelio Ladrón de Guevara Cortés.