

Neural Networks Principal Component Analysis for Estimating the Generative Multifactor Model of Returns under a Statistical Approach to the Arbitrage Pricing Theory: Evidence from the Mexican Stock Exchange

Rogelio Ladrón de Guevara Cortés¹, Salvador Torra Porras², Enric Monte Moreno³

¹ Universidad Veracruzana,
Institute for Research and Graduate Studies in Administrative Sciences (IIESCA),
Mexico

² University of Barcelona. Faculty of Economics and Business,
Department of Econometrics, Statistics and Applied Economy,
Spain

³ 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. A nonlinear principal component analysis (NLPCA) represents an extension of the standard principal component analysis (PCA) that overcomes the limitation of the PCA's assumption about the linearity of the model. The NLPCA belongs to the family of nonlinear versions of dimension reduction or the extraction techniques of underlying features, including nonlinear factor analysis and nonlinear independent component analysis, where the principal components are generalized from straight lines to curves. The NLPCA can be achieved via an artificial neural network specification where the PCA classic model is generalized to a nonlinear mode, namely, Neural Networks Principal Component Analysis (NNPCA). In order to extract a set of nonlinear underlying systematic risk factors, we estimate the generative multifactor model of returns in a statistical version of the Arbitrage Pricing Theory (APT), in the context of the Mexican Stock Exchange. We used an auto-associative multilayer perceptron neural network or autoencoder, where the 'bottleneck' layer represented the nonlinear principal components, or in our context, the scores of the underlying factors of systematic risk. This neural network represents a powerful technique capable of performing a nonlinear transformation of the observed

variables into the nonlinear principal components, and to execute a nonlinear mapping that reproduces the original variables. We propose a network architecture capable of generating a loading matrix that enables us to make a first approach to the interpretation of the extracted latent risk factors. In addition, we used a two stage methodology for the econometric contrast of the APT involving first, a simultaneous estimation of the system of equations via Seemingly Unrelated Regression (SUR), and secondly, a cross-section estimation via Ordinary Least Squared corrected by heteroskedasticity and autocorrelation by means of the Newey-West heteroskedasticity and autocorrelation consistent covariances estimates (HEC). The evidence found shows that the reproductions of the observed returns using the estimated components via NNPCA are suitable in almost all cases; nevertheless, the results in an econometric contrast lead us to a partial acceptance of the APT in the samples and periods studied.

Keywords. Extraction of underlying risk factors, nonlinear principal component analysis, arbitrage pricing theory, mexican stock exchange.

1 Introduction

The Principal Components Analysis (PCA) and Factor Analysis (FA) have been the classic techniques used for extracting the underlying systematic risk factors of the generative multifactor model of returns in the statistical approach to the Arbitrage Pricing Theory (APT). Both techniques make a strong assumption about the multivariate Gaussianity of the observed variables; however, real life data sets, especially financial time series, are not normally distributed neither univariate nor multivariate, and this causes the application of a PCA or a FA to yield unreliable results.

A solution to this problem is to extract the components by means of the Independent Component Analysis (ICA), which is capable of extracting statistically independent components from a set of non-Gaussian data. In addition, the underlying risk factors extracted by an ICA represent better estimations than those extracted by a PCA and a FA, because the first are statistically independent, whereas the latter are only linearly uncorrelated.

Nevertheless, both the PCA-FA and the ICA make another strong assumption: the linearity of the model. In the present research we use a novel extraction technique which deals with the nonlinearity problem: the Nonlinear Principal Components Analysis (NLPCA). This technique has been used in many fields of science as a dimensionality reduction or as a feature extraction technique¹.

For example, in [25] authors use a NLPCA to detect nonlinearities, extract features and classify spectral data from a set of stars, showing that the nonlinear principal components perform better than a standard PCA. They also apply it in the physiology field, analyzing data from electromyographic recordings of muscle activities and obtained similar results. In biochemistry and bioinformatics, in [26, 27, 21, 23] the authors applied a NLPC in order to analyze molecular

data from metabolite levels of a plant and from the reproductive cycle of a parasite.

Their findings demonstrate that the nonlinear components extracted by a NLPCA are more suitable for interpreting this kind of large multi-dimensional biological data as well.

Other fields of applications where there is an extensive list of studies are for instance: in Oceanography and Atmospheric Sciences, for extracting features from different atmospheric phenomena; in chemical and industrial engineering, for detecting faults in nonlinear industrial and chemical separation processes; in psychology, for dealing with nonlinear relationships applied to categorical data; and in robotics, for characterizing humanoid motion and for transferring human skills to robots.

In the field of finance, the application of NLPCA has been little developed. In [6] the authors used a NLPCA to determine the nonlinear principal components driving the variations of the implied volatility smile derived from FTSE-100 stock index options; in [18] a NLPCA is employed for bankruptcy prediction in banks, and in [30] it is used to analyze and predict the trend of withdrawals from an employment time guarantee fund.

On the other hand, some works have used related techniques to extract nonlinear components in the field of finance, e.g., [8] and [14], where the authors employ a Kernel PCA (KPCA) and a Curvilinear Component Analysis (CLCA), respectively, to reduce the dimension from a set of technical analysis indicators that they use for predicting stock prices and a market index. In addition, in [28] the authors used a KPCA to extract features from a set of stock prices with predictive purposes as well.

Applications in other related areas such as economics and business are limited too. In [13] the authors used NLPCA to evaluate the nonlinear relationship between budget rules and fiscal performance, and in [5] it is used as a dimensionality reduction technique to measure the perception of consumers about the quality of services.

in the feature extraction, the concern is for identifying unique and meaningful components or factors representing the main characteristics of the variables.

¹ The main difference between these two approaches is the required aim of the components or factors extracted. Whereas in the dimensional reduction the only interest is in achieving a smaller dimension of usually noise-free variables;

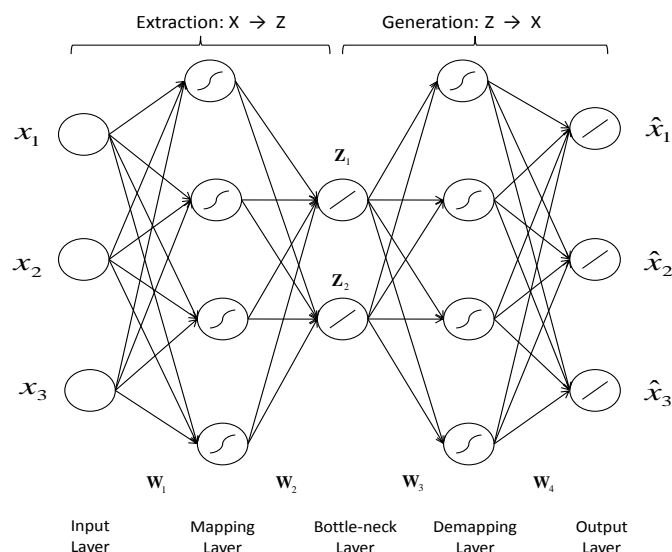


Fig. 1. Auto-associative multilayer perceptron neural network or autoencoder. Based on [23 and 16]

As far as we are concerned, there is neither any reference using NLPCA to extract the underlying systematic risk factors affecting the returns on equities in the stock markets, nor any study using NLPCA applied to Mexico; consequently, the main objective of this research is to fill this gap in financial literature².

The structure of this paper is as follows: Section 2 presents a brief review of the NLPCA, Section 3 explains the empirical study and Section 4 draws the conclusions. Finally, the last section presents the references consulted in this research.

2 Nonlinear Principal Component Analysis

The objective of a NLPCA is to extract nonlinear components from a data set.

The NLPCA represents a nonlinear generalization of the standard PCA, where the

² Nevertheless, there are some related studies carrying out other types of PCA, in the context of Ibero-America, to other fields of science that it is important to cite, such as: in [29] where the Dynamic Principal Component Analysis is compared to the Diagnostic Observers for fault detections in a heat exchanger; and in [2] where the authors used the Class-Conditional Probabilistic Component Analysis for gender recognition.

estimated principal components are taken from straight lines to curves, capable of handling and of discovering nonlinear relationships among variables and between components and variables, in other words, the subspace produced in the original data space is curved.

On the other hand, continuous financial time series, such as returns on equities, might present a nonlinear behavior, implying that possibly they might be better explained by curved lines rather than straight lines.

The relationship between the underlying systematic risk factors and the returns on equities might be nonlinear, too; thus, it could be better explained by a nonlinear model as well³.

2.1 Auto-Associative Neural Network

In this study we will focus on one approach to perform NLPCA based on artificial neural networks (ANN)⁴.

³ NLPCA belongs to the family of nonlinear versions of dimensionality reduction or feature extraction techniques, including Nonlinear Factor Analysis (NLFA) and Nonlinear Independent Component Analysis (NLICA). In addition, another related nonlinear approach in structural analysis is the Nonlinear Partial Least Squares (NLPLS).

⁴ Other methods to extract nonlinear components are: the Locally Linear Embedding (LLE), the Isometric Feature

This approach, known as Neural Networks Principal Component Analysis (NNPCA) or Principal Component Neural Networks (PCNN), is commonly performed via an auto-associative neural network architecture named autoencoder, replicator network, bottleneck or sandglass type network⁵. This neural network (NN) is a multilayer perceptron⁶, where the output layer of the network is required to be identical to the input layer (identity mapping) by minimizing the square error:

$$\|x - \hat{x}\|^2. \quad (1)$$

In the middle of the network there is a layer (bottleneck) where the reduction of dimension is done and represents the values of the principal components or scores. Figure 1 shows a diagram of this kind of NN.

The first part of the process is the extraction of the principal components (third layer) from the original data (first layer). The NN estimates a first matrix of weights to generate the second hidden layer, which will represent a previous layer before the one of nonlinear principal components (NLPC); then, the NN estimates a second matrix of weights, which will generate the third layer or principal components (Z).

The second part of the process is the reconstruction of the variables from the NLPC. The NN computes a third matrix of weights to produce a fourth hidden layer as a previous step to the reconstructed variables, which will be used together with to the fourth matrix of weights, in order to reproduce the original variables. Actually, the second and fourth hidden layers are the ones that perform the nonlinear mapping.

The formal expressions of the extraction and generation functions are:

Extraction function:

$$z = \phi_{extr} = (w, x) = \mathbf{W}_2 g(\mathbf{W}_1 x), \quad (2)$$

Generation function:

$$\hat{x} = \phi_{gen} = (w, z) = \mathbf{W}_4 g(\mathbf{W}_3 z), \quad (3)$$

Mapping (Isomap), the Principal Curves, the Self Organizing Maps (SOM), the Kernel PCA (KPCA), the Curvilinear Component Analysis (CLCA), and the Quantum-Inspired Evolutionary Algorithm (QIEA).

⁵ Another approach used for estimating the NLPCA based on NN is the input training network (IT-net). For details see [16].

where z represents the scores or principal components; \mathbf{W}_1 and \mathbf{W}_2 , the matrices of weights in the extraction process; \hat{x} , the reconstructed variables; \mathbf{W}_3 and \mathbf{W}_4 , the matrices of weights in the generation process; and g , the nonlinearity performing the nonlinear transformation, usually a tangent sigmoid function.

There are several architectures for the auto-associative neural network approach, such as: the standard, the hierarchical, the circular and the inverse model, and all of them can be used in combination. The standard NLPCA is the naive model, where both of the extraction and generation processes are included and no additional constraint, regarding the order of components, is imposed. The use of this version is recommended for non-periodic or non-cyclic data when the main interest is only the reduction of the dimensionality and not the extraction of meaningful features. In the hierarchical NLPCA, the order of the nonlinear components is enforced to respect the hierarchical ranking obtained in the linear PCA, thus yielding more meaningful features for the analysis. The circular version allows the extraction of circular components which describe a closed curve, instead of a standard curve with an open interval, more suitable for periodic or cyclic phenomena. Finally, the inverse definition only models the generation process.

This version is more efficient since we only train the second part of the neural network and not the two processes. It produces results more suited for real processes, since it models the natural process generating the observed data. In addition, it allows dealing with missing data because it does not need the sample data as an input. All the former extensions can be used in combination or separately⁷.

2.2 Dealing with Nonlinearity

In many studies a NLPCA has been used as a successful alternative to deal with the nonlinear

⁶ For details on multilayer perceptron neural networks, and in general, on foundation of neural networks, see [3].

⁷ For details about the different variants of the autoencoder NN, see [25, 26, 27, 21, 23].

relations among variables existent in different kinds of real data. Nevertheless, the use of NLPCA can be justified under a different perspective independently of the linear or nonlinear relation among the data set. Whereas PCA, FA and ICA represent linear models, a NLPCA has the attribute of being a nonlinear system⁸. In other words, PCA, FA and ICA express the variables in the model as linear combinations, while a NLPCA does it as a nonlinear mixing.

In NLPCA performed via an autoencoder neural network, the nonlinear hidden layers enable, first, a nonlinear mapping from the observed variables in order to estimate the principal components, and then another nonlinear transformation (demapping) from the estimated components in order to approximate the reconstructed variables. As a nonlinear system characterized by the non-proportionality between its inputs and outputs, a NLPCA will produce different insights of the studied phenomena. Particularly in the finance field, it could be assumed that simple variations in the underlying systematic risk factors may generate complex effects in the returns on equities; i.e., the relation between the stock returns and the underlying systematic risk factors may be nonlinear.

2.3 Estimation of the Parameters of the Model

The generation function, gives the inverse function, from a set in the latent space z , as shown in equation (3) above. In order to estimate the parameters of the model \mathbf{W}_3 and \mathbf{W}_4 that allow for the estimation of \hat{x} , we will optimize a global cost function that simultaneously yields, a projection from the initial values x into the latent space and then from the latent space x estimates a new reconstruction \hat{x} . According to [21] this is done by minimizing the RMSE as follows:

$$E(w) = \frac{1}{N} \sum_n \|x_k^n - W_4 g_2(W_3 W_2 g_1(W_1 x_k^n))\|^2 \quad (4)$$

⁸ As stated in [21]: "Linear models can be expressed as a (weighted) sum of their individual parts (factors ...). Nonlinear models, by contrast, cannot simply be expressed by a sum. More precisely, the linear transformation ... of a linear model is given by a linear function. A function $f(x)$ is termed linear

In the case of implementation done in this paper, the parameters w_{ij} corresponding to each of the matrices $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4$ were estimated by means of gradient search.

Following [26] in order to compute each of the weight values w_{ij} the above equation can be expressed more specifically as:

$$E(w, z) = \frac{1}{dN} \sum_n \sum_i^d [x_i^n - \sum_j^h w_{ij} g(\sum_i^m w_{jk} z_k^n)]^2, \quad (5)$$

and the partial derivatives are the following (according to the Matlab® notation, and including the bias term corresponding to $j = 0$):

$$\frac{\partial E}{\partial w_{jk}} = \sum_n \sigma_i^n g(a_j^n); \quad \sigma_i^n = \hat{x}_i^n - x_i^n, \quad (6)$$

$$\frac{\partial E}{\partial w_{jk}} = \sum_n \sigma_j^n z_k^n; \quad \sigma_j^n = g'(a_j^n) \sum_i w_{ij} \sigma_i^n, \quad (7)$$

$$\frac{\partial E}{\partial z_k^n} = \sigma_k^n; \quad \sigma_k^n = \sum_j w_{kj} \sigma_j^n, \quad (8)$$

The above partial derivatives are used for updating the estimation of the weights w_{ij} iteratively as:

$$w_{ij} \leftarrow w_{ij} - \mu \frac{\partial E}{\partial w_{ij}}. \quad (9)$$

The selection of the μ mu parameter is done heuristically, taking into account convergence considerations.

3 Empirical Study

3.1 The Data

In conformity with the availability of information, and for the sake of a comparison to former studies⁹, the sample used in this research contains the log returns of the stocks that have been part of the main index of the Mexican Stock

when it satisfies both properties: additivity $(x + y) = f(x) + f(y)$ and homogeneity $(ax) = af(x)$, otherwise it is a more complex nonlinear function."

⁹ See [10, 11].

Table 1. Stocks used in this study

No.	TICKER	Name of the Company	Industrial Sector
1	ALFAA	Grupo Alfa	Holding
2	ARA*	Consortio Ara	Construction: Housing
3	BIMBOA	Grupo Bimbo	Food processing
4	CEMEXCP	Cemex	Cement
5	CIEB	Corporación Interamericana de Entretenimiento	Holding
6	COMERUBC	Controladora Comercial Mexicana	Commerce: retailing and wholesale
7	CONTAL*	Grupo Continental	Food and beverage processing
8	ELEKTRA*	Grupo Elektra	Commercial firms
9	FEMSAUBD	Fomento Económico Mexicano	Beer
10	GCARSOA1	Grupo Carso	Holding
11	GEOB	Corporación GEO	Construction: Housing
12	GFINBURO	Grupo Financiero Inbursa	Financial services
13	GFNORTEO	Grupo Financiero Banorte	Financial services
14	GMODELOC	Grupo Modelo	Food, tobacco and beverages
15	KIMBERA	Kimberly-Clark de México	Cellulose and paper
16	PE&OLES*	Industrias Peñoles	Ferrous minerals
17	SORIANAB	Organización Soriana	Commerce: retailing and wholesale
18	TELECOA1	Carso Global Telecom	Communications
19	TELMEXL	Teléfonos de México	Communications
20	TLEVICPO	Grupo Televisa	Communications
21	TVAZTCPO	TV Azteca	Communications
22	WALMEXV	Wal-Mart de México	Commerce: retailing and wholesale

Exchange, the Price and Quotations Index (IPC), during all the periods considered.

Because of their importance in the Mexican Economy and their characteristics of liquidity and market value, these companies can be considered as representative of the Mexican stock market. Table 1 shows the names and sectors of these shares.

We carried out our study on four different databases structured as follows: two databases of 20 stocks ranging from July 7, 2000 to January 27, 2006, one expressed in weekly returns

¹⁰ The two stocks not included in the daily databases are: CEMEXCP and KIMBERA. The interest rates considered as the riskless interest rate were the average weekly and daily

(DBWR) and the other in returns in excess of the riskless interest rate (DBWE); and two databases of 22 stocks running through the period from July 3, 2000 to January 31, 2006, the first expressed in returns (DBDR) and the latter in excesses (DBDE)¹⁰.

The period analyzed in this study (2000-2006) was considered according to the following criteria:

a) This article represents the third part of a major research, where we are testing different techniques for extracting the underlying

funding interest rates using government securities, published by the Bank of Mexico, <http://www.banxico.org>

systematic risk factors in the context of the Mexican Stock Exchange. The Neural Networks Principal Component Analysis represents the third approach we have used to perform that extraction, under a statistical approach to the latent systematic risk factors.

- b) The technique used in this article, and the other techniques utilized in the previous stages of our research, have an explanatory and a predictive character. First, we are carrying out the explanatory approach, which make us divide our dataset in two blocks: one for explanation or training and another for prediction; i.e., the first period is used for the explanation or training of the model, and the second one, will be used for testing the predictive power of the generative model of returns estimated. Consequently, the data from 2000 to 2006 were used to extract the generative underlying structure of returns, which explains the behavior of the returns of the training period. This estimation will help us in the next stage of prediction, where the model will be tested in subsequent periods of time (from 2006 on).
- c) The other techniques that we have employed in our research are the Independent Component Analysis, Factor Analysis and Principal Component Analysis; our objective is to be able to compare the results of the four techniques, concerning both their explanatory and predictive power. Therefore, we are using the same training and prediction periods for the four of them.
- d) Additionally, another reason for using this period of the dataset, is to be able to compare in further studies, the effects of the 2008-2009 financial crisis in the estimation of the

underlying structure of systematic risk, by means of the extraction of the generative model of returns during the crisis and the post-crisis periods, using the four techniques¹¹.

3.2 Methodology and Results

3.2.1 Extraction of Underlying Systematic Risk Factors Via NNPCA

The Arbitrage Pricing Theory (APT) assumes the following generative multifactor model of returns¹²:

$$R = E(R_i) + \beta_{1i} \cdot F_{1t} + \beta_{2i} \cdot F_{2t} + \dots + \beta_{ji} \cdot F_{jt} + \varepsilon_{it}. \quad (10)$$

From the statistical approach, neither the factors nor their sensitivities are given¹³ and we must estimate them simultaneously by means of statistical or feature extraction techniques such as, in this case, the NNPCA. Although the NNPCA is capable of extracting the scores of the components (the F_s), it is very difficult or even impossible to obtain a single matrix containing the equivalent to the sensitivities for each factor (betas), because there are two matrices of weights and a nonlinear transformation involved in the process of reproducing the variables¹⁴. Consequently, we used the NNPCA for extracting only the scores of the underlying systematic risk (the F_s) in the expression 10.

In order to estimate the NNPCA model, we used its hierarchical extension (h-NLPCA) performed by an auto-associative neural network, which respects the ranking of the principal components in the linear PCA¹⁵. This hierarchy implies the fulfillment of two important properties: scalability and stability. Scalability means that the first n components must explain, as much as

¹¹ This study only focuses in the estimation of the explanatory model using NNPCA. The estimation of the explanatory models using the other referred techniques, the testing of the prediction power of the estimated models and the comparison of the results in the crisis and post-crisis periods are out of the scope of this article, and represent other stages of the research conducted by the authors of the present document.

¹² Where, β_{ji} represents the sensitivity of equity i to factor j , F_{jt} the value of the systematic risk factor j in time t common for all the stocks, and ε_{it} the idiosyncratic risk affecting only the equity i .

¹³ The macroeconomic approach assumes the factors as given and estimates the sensitivities; whereas the fundamental, sector, and technical approaches assume the sensitivities as given and estimate the factors [31,1].

¹⁴ The analogous situation occurs in the process of extraction, where there are two matrices of weights and a nonlinear transformation involved in the extraction of the components process.

¹⁵We used the Matlab® code created by [22] to perform the NNPCA estimation, available at <http://www.nlpca.org/matlab>.

possible, the variance in the n -dimensional subspace. Stability denotes that the i -th component of an n -dimensional solution must be identical to the i -th component of an m -dimensional result, where $n \neq m$.

According to [25], the hierarchy constraints are based on searching in the original data space for the smallest mean square reconstruction error while using the first i -th components according to the following expression:

$$E = \frac{1}{dN} \sum_n \sum_k^d (x_k^n - \hat{x}_k^n), \quad (11)$$

where, x and \hat{x} represent the observed and reproduced data respectively; N , the number of samples; and d , the dimensionality. The hierarchical error function extended to k components, with $k < d$, implies minimizing:

$$E_H = E_1 + E_{1,2} + E_{1,2,3} + \dots + E_{1,2,3,\dots,k}. \quad (12)$$

Therefore, the h-NLPCA can be interpreted as we look for a k -dimensional subspace of minimal mean square error (MSE), so that the $(k-1)$ -dimensional subspace is also of MSE. Consequently, all the dimensional subspaces $1, k$, are of minimal MSE and represent their dimensionality in the best way¹⁶. For the sake of a comparison with our former studies, we estimated 8 different Neural Networks (NNs) to extract from 2 to 9 nonlinear principal components in each database.

In order to generate a loading matrix that make possible a first attempt to interpret the latent risk factor extracted, we used a five-layer architecture with 20 neurons in the input layer for the weekly databases and 22 for the daily ones, from 2 to 9¹⁷ in the mapping layer, the bottleneck layer, and the demapping layer and, finally, 20 and 22 in the output layer, respectively. In terms of the NN notation, the architectures used were: [20:2-9:2-9:2-9:20] and [22:2-9:2-9:2-9:22].

Concerning the nonlinear transferring functions, following the recommendations of [16] for an autoencoder NN in order to perform the

NLPCA, we used a tangent sigmoid function from layer one to layer two and from layer three to layer four; and a linear function from layer two to layer three and from layer four to layer five.

Using the Matlab® code by [22] for the performance of the NLPCA on our four databases, we obtained the scores of the principal components hierarchically ordered, the four matrices of weights and the reproduced variables. We emphasize that the objective of such estimation is to achieve a nonlinear transformation, first, from the observed variables to the principal components, and then to realize another nonlinear transformation capable of reproducing the observed variables from the extracted components.

The results in the reconstruction of the observed returns or excesses were suitable for all the stocks in the four databases; this implies that the estimation of the generative multifactor model in the statistical approach to the APT, performed by NLPCA, was successful¹⁸.

The only problem detected was in the reproduction of some observations in a few stocks presenting very high levels of volatility, where the reconstruction was not able to reach all the peaks completely. Nevertheless, if we add more components to the extraction, the reproduction of all the series improves greatly, covering almost all the peaks of high volatility¹⁹.

For reasons of space saving, in Figures 2 we only show the lines plots of the observed and reproduced returns and excesses from the first 5 stocks in each database of the experiment, where we extracted nine components. We can observe that the reconstruction is suitable in nearly all cases, except for the observations regarding very high volatility, as stated above.

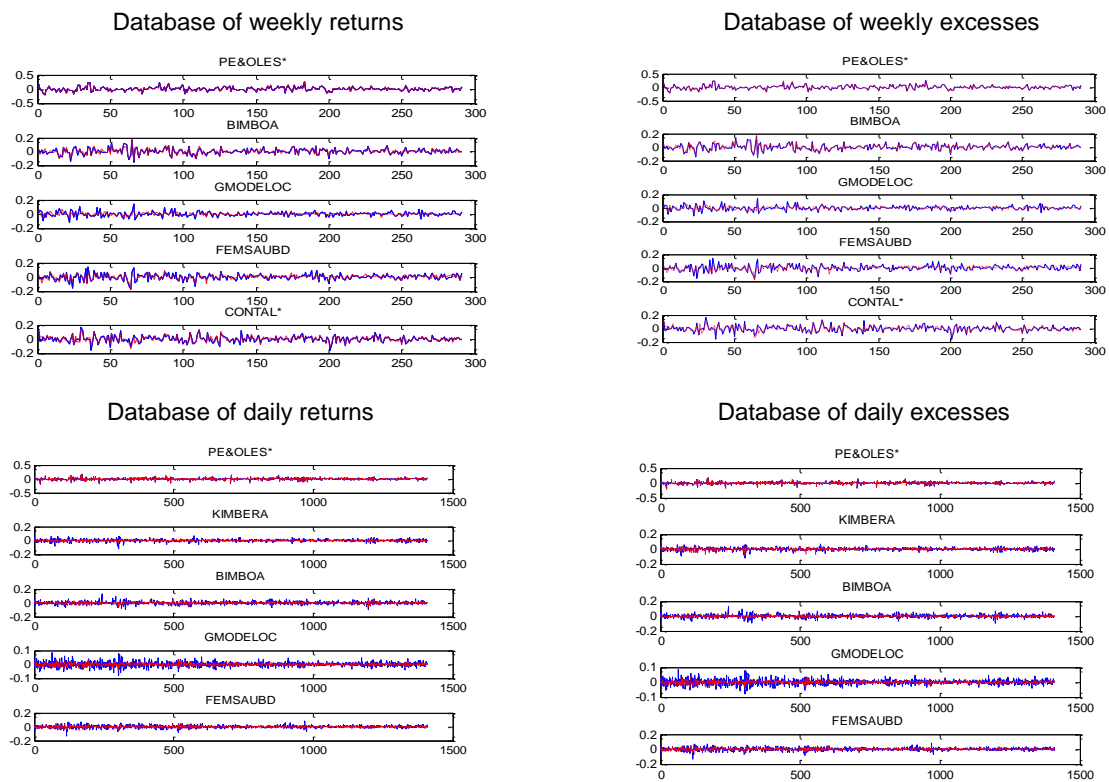
In addition, for visualization purposes in Figures 3, we present the plots generated by the software used for the extraction, where the first three principal components of the NLPCA are plotted as a grid in the original data space.

¹⁶For details on the hierarchical error function, see [25].

¹⁷ Our test window in each database run from 2 to 9 non-linear components.

¹⁸ Obviously, the greater the number of components estimated, the better the reproduction capacity of the model.

¹⁹ These experiments are not reported because we only focused on the range of estimation from 2 to 9 components. In spite of these findings, further research about this loss of volatility in the reconstruction might be done, as an attempt to discover why these components fail in picking up the risk.



Note: Logarithmic returns of the first five stocks observed in each database and their respective reconstructions using the estimated NNPCA model. Stock symbols of the stocks presented appear above each line plots.

Fig. 2. Line plots of the observed and reproduced stocks

In this case the grids represent the new coordinates of the component space, thus giving a nonlinear or curved description of the data.

Although it is not completely conclusive, the four plots show that the data could be described sufficiently well by nonlinear behaviors.

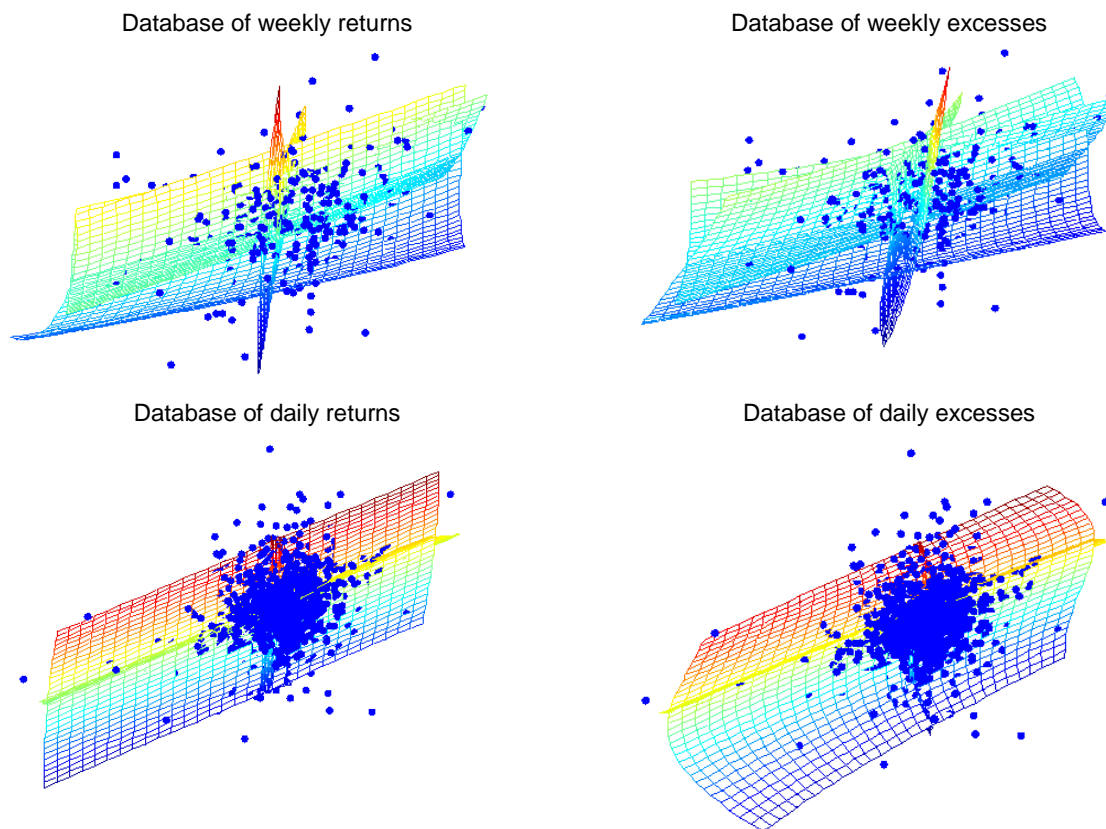
3.2.2 Interpretation of the Extracted Factors

Although this study is mainly focused on the extraction process of systematic risk factors of the Mexican Stock Exchange, but not on the risk attribution stage of statistical approach to the Arbitrage Pricing Theory, in this section we will just make a first attempt to propose an interpretation of the meaning or nature of the systematic risk factors extracted. We will follow an analogue methodology similar to the classical approach used when Principal Component Analysis (PCA) and Factor Analysis (FA) are used

to reduce dimensionality or to extract features from a multifactor dataset.

This approach is based on the use of the factor loading matrix estimated in the extraction process in order to identify the loading of each variable in each component or factor; high factor loadings in absolute terms indicate a strong relation between the variables and the factor. In our context, the factors will be saturated with loadings of one stock or a group of stocks that may help us in the identification of those factors with certain economic sectors, as a first approach to the interpretation of each component or factor.

In the case of NNPCA, that factor loading matrix is not clearly defined, since the demixing process involves the combined effect of two loading matrices (\mathbf{W}_1 and \mathbf{W}_2) and a nonlinear function of transference; however, in order to use one of these matrices as an analogue one to



Note: The first three principal components of NNPCA plotted as grid in the original data space. The grids represent the new coordinates in the space of the components and give a nonlinear or curved description of the data.

Fig. 3. Nonlinear PCA plot. Nine components estimated

those used in techniques such as PCA and ICA as a first approach to give meaning to the extracted factors, we can argue the following, considering the role that each matrix plays in the demixing process.

Following the network architecture displayed in Figure 1. Matrix \mathbf{W}_1 makes a projection into the space where we have an internal representation in the form of the hidden units, thus, it would be equivalent to a mixing matrix such as those used in PCA and FA. In other words, from a structural point of view NNPCA makes a non-linear transformation given by \mathbf{W}_1 . For that effect, it is necessary to subtract the medium value by means of the bias involved in the estimation and to scale the inputs somehow, so that the nonlinearity compresses the margin properly.

This makes the function of the first layer of the network to be different to that of other methods such as PCA and FA. On the other hand, matrix \mathbf{W}_2 makes a dimensionality change of the representation given the output of the first layer.

Its function is to make a lineal transformation to rotate and scale the output, in such a way that the intermediate representation could be transformed by the second part of the network.

Furthermore, from a structural standpoint, the product $(\mathbf{W}_1 * \mathbf{x})$ in expression 1, generates the representation that will pass through the nonlinearity later.

The function of the nonlinearity is to make a compression of the space in order to make easy the function of the posterior part of the neural network easy. From this standpoint, the projection

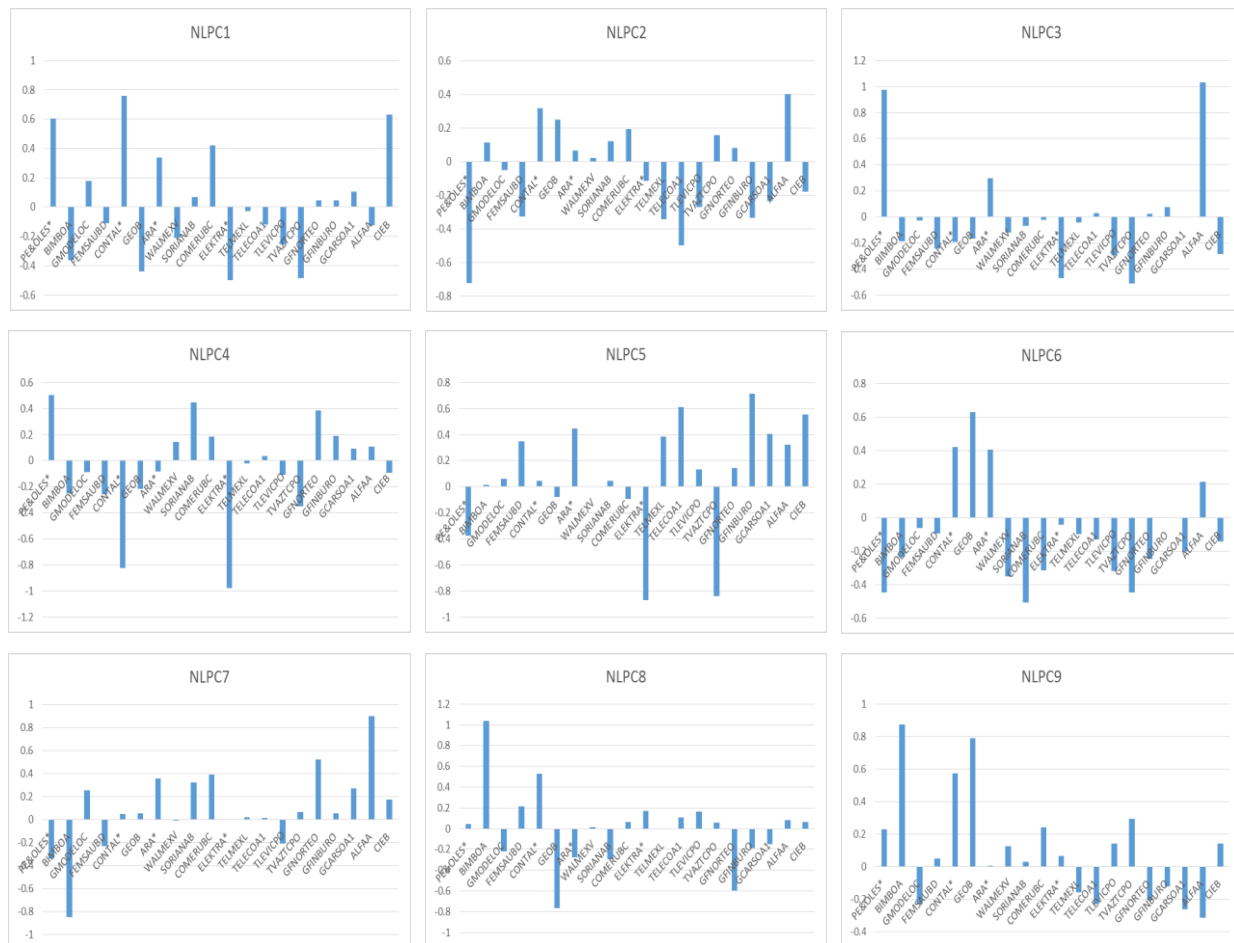


Fig. 4. Loadings matrices plots for interpretation of extracted factors. Database of weekly returns Nine components extracted

form given by $(W_1 \cdot x)$ informs about the intermediate representation of the information and it could be compared with the latent factors estimated by PCA and FA; although it is important to remark that they are different things since they are obtained through different criteria.

According to the above stated, in this research we will use matrix W_1 as a loading matrix to propose preliminary meanings for the extracted latent factors.

In the interest of saving space, we only present the loading matrices plots from the database of weekly returns that belong to the experiment where we extracted nine underlying factors, nevertheless, this kind of plots were developed for

all the cases under the same methodology. Figure 4 presents these results.

Additionally, we constructed some tables summarizing the results derived from the analysis of the factor loading matrices and plots, where we propose a certain economic sector that may be related to each factor. We grouped together the stocks with the highest loading in each factor according to the official classification of the economic sectors used in the Mexican Stock Exchange.

Table 2 presents this summary. There is not a clear interpretation of the factors using the matrix W_1 ; however, we uncover that in this case, the most of the factors are formed by a mixture of

Table 2. Summary of the Interpretation of the Non-linear Principal Components

Database of Weekly Returns		Database of Weekly Excesses	
NLPC1	Beverages and Leisure / mining sectors factor.	NLPC1	Mining / Food products and beverages, Consumer staples and Communication media sectors factor.
NLPC2	Mining and Telecommunications / Holding sectors factor.	NLPC2	Mining / House building sectors factor.
NLPC3	Holding / Mining sectors factor.	NLPC3	House building, Mining and Holdings sectors factor.
NLPC4	Home Furnishing and Beverages sectors factor.	NLPC4	Beverages, Leisure and Home furnishing sectors factor.
NLPC5	Salinas Group Factor.	NLPC5	Consume sector factor.
NLPC6	House building and Beverages / Consumer staples, Communication media and Mining sector factors.	NLPC6	Construction sector factor (Geo Factor).
NLPC7	Holdings / Food products sector factors.	NLPC7	Financial and House building /Consumer staples sectors factors.
NLPC8	Food products / Construction sector factors.	NLPC8	Food and beverages sector factor.
NLPC9	Food products, Beverages and Construction sector factors.	NLPC9	House building, communication media and consumer staples sector factor.
Database of Daily Returns		Database of Daily Excesses	
NLPC1	Construction sector factor (Geo factor).	NLPC1	Salinas Group / Mining sector factor.
NLPC2	Mining sector factor (Peñoles factor).	NLPC2	Beverages / Home furnishing and Financial services sectors factor.
NLPC3	Consumer staples, Financial services, Home furnishing and Mining sector factors.	NLPC3	Salinas Group, Holdings and Mining / Leisure sectors factor.
NLPC4	Communication media and Beverage sector factor.	NLPC4	Holdings / Leisure sectors factors.
NLPC5	Beverages and mining / Home furnishing and house building sectors factor.	NLPC5	Beverages and House building / Mining sector factors.
NLPC6	Beverages, Communication media, House building and Home furnishing sectors factor.	NLPC6	House building and Holdings / Leisure sectors factor.
NLPC7	Leisure and Financial services sectors / Salinas Group factor.	NLPC7	Communication media / Financial services sectors factor.
NLPC8	House building and Holdings / Home furnishing and Consumer staples sector factor.	NLPC8	Mining sector factor (Peñoles factor).
NLPC9	Holdings and House building / Mining and Home furnishing sector factors.	NLPC9	Mining and Beverages sector factor.

stocks from different industrial sectors instead of a combination of shares from the same sector.

In other words, excluding some factors that we could identify clearly; i.e.: number five (Salinas Group factor), in database of weekly returns;

number five (Consumer sector factor), number six (Construction sector factor or GEO factor) and number eight (Food and Beverage sector factor), in database of weekly excesses; number one (Construction sector factor or Geo factor) and

number two (Mining factor or Peñoles factor), in a database of daily returns; and finally, number eight (Mining factor or Peñoles factor), in database of daily excesses; the rest of the factors represent a combination of sectors that in many cases have opposite signs.

In addition, we can distinguish the strong and constant contribution of some sectors or stocks in many factors in the four databases; e.g., mining sector with PEÑÓLES (DBWR:4 + DBWE:3 + DBDR:4 + DBDE:5 = 16), beverage sector with CONTAL (DBWR:4 + DBWE:3 + DBDR:3 + DBDE:3 = 13), construction sector with GEO (DBWR:3 + DBWE:5 + DBDR:5 = 13), home furnishing sector with ELEKTRA (DBDR:6 + DBDE:3 = 9), holding sector with ALFA (DBWR:3 + DBDE:3 = 6), food products sector with BIMBO (DBWR:3), consumer staples with WALMEX (DBWE: 3) and SORIANA (DBWE:3), communication media sector with TVAZTECA (DBDR:3) and leisure sector with CIEB (DBDE:3).

In this case, none of the components in any database is clearly related to market factor. Likewise, there is not a homogeneous interpretation of the factors in all the databases. Nevertheless, there are two factors that could have the same interpretation in the different databases but are ranked in different order; e.g., the mining and the construction factors, as can be observed in the referred table.

3.2.3 Econometric Contrast

As a complement for our research, we carried out an econometric contrast of the APT, using the underlying systematic risk factors extracted via the NNPCA, in order to test its validity as a suitable pricing model for the sample and periods considered. This methodology of contrast represents only a first approach to the econometric validation of the APT using NNPCA, so the result should be viewed in that light.

After applying the second assumption of the APT (the principle of arbitrage) to its first assumption (the generative multifactor model of returns) of expression 10, we get the APT fundamental pricing equation²⁰:

²⁰ A clear analytic development about the integration of the two assumptions which yields the APT pricing equation can be seen in [1].

$$E(R_i) = \lambda_0 + \lambda_1 \cdot \beta_{1i} + \lambda_2 \cdot \beta_{2i} + \dots + \lambda_k \cdot \beta_{ki}, \quad (13)$$

where the betas are the sensitivities to the systematic risk factors and the lambdas are the risk premium paid by the market for being exposed to each class of systematic risk.

The former equation can be tested by means of an average cross-section methodology for estimating the ordinary least squares (OLS) coefficients of the following regression model:

$$\bar{R} = \lambda_0 + \lambda_1 \cdot \beta_{1i} + \lambda_2 \cdot \beta_{2i} + \dots + \lambda_k \cdot \beta_{ki} + \varepsilon. \quad (14)$$

The straight methodology for contrasting the APT under the statistical approach would directly use the loadings or betas estimated in expression 10 in the former regression model [5], since both factors and sensitivities are computed simultaneously by the extraction techniques usually employed [1].

Nevertheless, as remarked in [15, 17], this methodology could present some econometric problems such as heteroskedasticity and autocorrelation in the residuals, in addition to error in variables, which would yield inefficient OLS estimators with biased variances. Besides, the NNPCA estimation does not generate a single matrix equivalent to the loadings in the generative multifactor model of returns; hence, we cannot use this methodology. One possible solution to the foregoing problems is to employ a two-stage methodology widely used in the fundamental and macroeconomic approach to the APT, where in a first stage we estimate the betas to use in expression 14, then in a second stage we estimate the lambdas.

Following [4]²¹, in the first stage we estimated the betas to be used in expression 13, by regressing the factor scores obtained by the NNPCA 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 a seemingly unrelated regression (SUR) to estimate simultaneously the entire system of equations.

²¹In their work, the authors use principal component analysis to extract the underlying risk factors from a set of macroeconomic variables in the Spanish market.

Table 3. Summary of the Econometric Contrast

		λ_0	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	R^2	$\lambda_{sig} / \lambda_{tot}$	WALD	J_B
Database of weekly returns															
Model betas	with 2	•		•								9.45%	0.00%		•
Model betas	with 3	0.005078		0.01034	0.02173							51.89%	66.67%		•
Model betas	with 4	0.005582		0.00193	0.01002	•						48.58%	50.00%		•
Model betas	with 5	0.005411		0.00892	0.02423	•	0.00348					50.84%	60.00%		•
Model betas	with 6	0.004886		0.00378	0.00997	•	•	•				47.96%	33.33%		○
Model betas	with 7	0.005458		0.00362	0.01168	•	•	•				55.59%	28.57%		○
Model betas	with 8	0.005605		0.00303	0.02117	•	•	•	•	•		50.58%	25.00%		○
Model betas	with 9	0.005782		•	0.02016	•	•	•	•	•		46.35%	11.11%		○
Database of weekly excesses															
Model betas	with 2	•		•								6.61%	0.00%		○
Model betas	with 3	0.003488		0.00195	0.02129							47.35%	66.67%		•
Model betas	with 4	0.003945		0.00237	0.00481	•						49.04%	50.00%		•
Model betas	with 5	•		0.00505	0.03206	•	•					48.14%	40.00%		○
Model betas	with 6	•		0.00404	0.00882	•	0.00147	•				52.74%	50.00%		○
Model betas	with 7	•		0.00218	0.00650	•	0.00168	•				51.67%	42.86%		○
Model betas	with 8	•		0.00439	0.02272	•	•	•	•			53.0%	25.00%		○
Model betas	with 9	0.0433		0.00613	0.02391	•	•	•	•	0.00040		57.13%	33.33%		•
Database of daily returns															
Model betas	with 2	•		•								0.00%	0.00%		○
Model betas	with 3	0.00047		0.00113	0.00104	-						38.93%	66.67%		○
Model betas	with 4	•		0.00090	0.00184	-	•					38.11%	50.00%		○
Model betas	with 5	•		•	0.00229	-	•	•				44.15%	20.00%		○
Model betas	with 6	0.001226		•	0.00401	-	•	•	•			56.56%	16.67%		○
Model betas	with 7	•		•	0.00211	-	•	•	•			50.05%	14.29%		○
Model betas	with 8	•		•	0.00163	-	•	•	•	•		49.49%	12.50%		○
Model betas	with 9	•		•	0.00361	-	•	•	0.00058	•	•	61.79%	22.22%		○

Database of daily excesses										
Model with 2 betas	•	0.00046				1.36%	50.00%	○	○	
Model with 3 betas	•	0.00085	0.00162			41.19%	66.67%	○	○	
Model with 4 betas	0.000636	0.00043	0.00140			50.91%	50.00%	•	•	
Model with 5 betas	•	0.00080	0.00174	•		36.09%	40.00%	○	•	
Model with 6 betas	•	•	0.00402	•		48.02%	16.67%	○	○	
Model with 7 betas	•	•	0.00146	0.00065	•	44.49%	28.57%	○	○	
Model with 8 betas	•	•	0.00284	0.00069	0.00028	62.43%	37.50%	○	○	
Model with 9 betas	•	•	0.00281	•	•	•	55.92%	11.11%	○	○

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) λ_j : Estimated coefficients. $H_0: \lambda_j = 0$. 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) *Wald*: Wald's test for coefficient restrictions. Databases in returns: $H_0: \lambda_0 =$ Average riskless interest rate. Databases in excesses: $H_0: \lambda_0 = 0$. ○ = Not rejection of H_0 . The independent term is equal to its theoretical 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. $H_0 =$ Normality. ○ = Not rejection of H_0 . The residuals are normally distributed. • = Rejection of H_0 . The residuals are not normally distributed.

The results of the regressions in the four databases were very good, producing, in almost all cases, statistically significant parameters, high R^2 coefficients and statistics from the Durbin-Watson test of autocorrelation, all of which led us to the non-rejection of the null hypothesis of no-autocorrelation²².

Following [9]²³, in the second stage we estimated the lambdas in expression 14 by regressing the betas obtained in the first stage as a cross-section on the returns and excesses, using OLS. In order to avoid the econometric problems of heteroskedasticity and autocorrelation in the residuals of the model estimated through OLS, we used Ordinary Least Squared corrected by heteroskedasticity and autocorrelation by means of the Newey-West heteroskedasticity and autocorrelation consistent covariances 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 λ_0 ²⁴, and the

equality of the independent term to its theoretical value, i.e., the average returns, in the models expressed in returns:

$$\lambda_0 = \bar{R}_0, \quad (15)$$

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

$$\lambda_0 = 0. \quad (16)$$

We used Wald's test to confirm these equalities.

In Table 3, we present a summary of the results of the econometric contrast. In general, the results of the explanation power (R^2), the statistical significance of the multivariate test (F), and the residual test are very good in all the contrasted models, except in the cases where only two factors were extracted.

The univariate tests for the individual statistical significance of the parameters²⁵ priced from one to three factors different from λ_0 , thus, giving evidence in favor of the APT in 29 models²⁶. Nevertheless, only four models fulfilled both the

²² For reasons of saving space these results are not presented.

²³ In their study the authors used a factor analysis to extract the underlying risk factors from a set of returns on mutual funds in the Spanish market.

²⁴The ideal situation is that more than one parameter different from λ_0 be statistically significant, since the APT assumes that there are multiple underlying risk factors in the economy affecting the returns on equities and not only one.

²⁵ Statistic t .

²⁶ The total number of tested models was 32.

statistical significance and the equality of the independent term to its theoretic value, in addition to the fulfilment of the requirements imposed by the residual test.

These two models were those expressed in weekly returns when six, seven and eight factors were extracted; and the one expressed in daily returns when three components were estimated. Moreover, there are twelve other models which fulfil all the conditions for accepting the APT as a pricing model, except for the statistical significance of the independent term, and eight models that fail only in the equality of the independent term to its theoretical value, which provides some additional evidence in favor of this asset-pricing model.

Making a cross-validation with the interpretation of the factors proposed in section 3.2.2, the meaning of the significant factors corresponding to the fully accepted models are the following. In the four models the statistical significant factors were number two and three.

Regarding the database of weekly returns, factor number two contrast the Mining and Telecommunications sectors to the Holding sector; and number three, counter the Holding sector to the Mining one. Concerning the model in the database of daily returns, factor number two was related to the Mining sector factor (Peños Factor); and number three correspond to a factor that mixes stocks of the Consumer staples, Financial services, Home furnishing and Mining sectors. Interestingly, datasets expressed in excesses did not produce any fully accepted model.

Further research will be needed regarding this issue, as well as the significance of the undersized values and signs of the estimated individual parameters. To summarize, for the sample and periods considered, we can accept only partially the validity of the NNPCA-APT as a pricing model explaining the average returns (and returns in excesses) on equities of the Mexican Stock Exchange. On the other hand, the evidence showed that the APT is sensitive to the number of

factors extracted and to the periodicity and expression of the models.

4 Conclusions

The theoretical attributes of the NLPCA present desirable features when we extract the underlying systematic factors via this alternative technique, since they represent nonlinearly uncorrelated factors and not only linearly uncorrelated ones. The NNLPCA performed via NNPCA is capable of uncovering both, linear and nonlinear correlations, while PCA identifies, for example, only linear correlations. In that sense, we may conclude that the factors obtained in this study represent a more desirable estimation of the underlying systematic risk factors under a statistical approach to the APT²⁷. In our case, we believe that the extracted factors should be better estimations²⁸, to be used in a statistical approach to the APT because: first, they represent factors that have eliminated both, linear and nonlinear correlations among variables, and second, they are the result of a nonlinear transformation, not only a linear mapping, which deals with any nonlinear effect of the systematic risk factors over the returns on equities.

In addition, it is important to point that the non-Gaussian nature of the financial data, make that the generally used techniques for extracting underlying risk factors, such as PCA or FA, may generate, not completely reliable estimations, which suggest that the estimation of the generative underlying multifactor model of returns on equities by means of NNPCA, could represent a more reliable option to this end. [See 11, 12, 17, 9, 4, 5].

We would like to remark that our main goal in this paper has been the estimation of the generative multifactor model of returns of the APT by means of the NNPCA, that is, the risk extraction stage of a statistical approach to the APT. Therefore, the interpretation of the components extracted represents only a first attempt to give meaning to the latent factors;

factors would suppose a better attribute of those extracted factors.

²⁸ Nevertheless this statement is object of academic discussion.

²⁷ More desirable in the sense that under the scope of the APT, in general, and the statistical approach, in particular, we look for obtaining risk factors as much independent or different as possible. In that sense having nonlinearly uncorrelated

however, further research will be needed about the risk attribution process of this statistical approach.

In the same way, the econometric contrast corresponds only to a first approach to the validation of the APT as a pricing model using the systematic risk factors estimated via this extraction technique; therefore its results should be seen under this perspective. For the moment, we could attribute the not completely satisfactory results of the econometric contrast to two possible reasons: a) The methodology used for the contrast might not be the most suitable for a statistical approach to the APT, and perhaps it would be necessary to use time series moving regressions to estimate the sensitivities to the risk factors or betas [17, 19] or mimicking portfolios as proxies of the underlying factors [15, 29]. b) The origin of the problem might not be in the first assumption of the APT, the generative multifactor model of returns, but in the second, the arbitrage absence principle [10]; aspect that we have not investigated yet. Further research would be needed concerning these two possible causes of the results in the econometric contrast.

Acknowledgments

The authors thank Matthias Scholz from the Fondazione Edmund Mach for the technical advice on some topics related to this investigation, and Cristina Urbano at Gaesco for the financial data provided.

References

1. **Amenc, N. & Le-Sourd, V. (2003).** *Portfolio Theory and Performance Analysis*. USA: Wiley.
2. **Bekios, J., Buenaposada, J., & Baumela, L. (2011).** Class-Conditional Probabilistic Principal Component Analysis: Application to Gender Recognition. *Computación y Sistemas*, Vol. 14, No. 4, pp. 383–391.
3. **Bishop, C.M. (1995).** *Neural networks pattern recognition*.
4. **Bruno, N., Medina, U., & Morini, S. (2002).** Contraste factorial del arbitrage pricing theory en el mercado bursátil español. *Análisis Financiero* Vol. 88, pp. 58–63.
5. **Gómez-Bezares, F., Madariaga, J., & Santibáñez, J. (1994).** *Valoración de acciones en la Bolsa Española*.
6. **Fan, K., O'Sullivan, C., Brabazon, A., & O'Neill, M. (2008).** Non-linear principal component analysis of the implied volatility smile using a quantum-inspired evolutionary algorithm. In **A. Brabazon & M. O'Neill (eds.)**, *Natural Computing in Computational Finance*, pp. 89–108. DOI: 10.1007/978-3-540-77477-8_6.
7. **Ferrari, P. & Salini, S. (2008).** *Measuring service quality: the opinion of Europeans about utilities*. FEE working paper, No. 36, pp. 1–29. DOI: 10.2139/ssrn.1123414.
8. **Ince, H. & Trafalis, T. (2004).** Kernel principal component analysis and support vector machines for stock price prediction. *IEEE Transactions*, Vol. 39, No. 6, pp. 2053–2058. DOI: 10.1109/IJCNN.2004.1380933.
9. **Jordán, L. & García, J. (2003).** Estimación y contraste del modelo APT en los fondos de inversión mobiliaria españoles. *Análisis financiero*, No. 89, pp. 22–35.
10. **Khan, M. & Sun, Y. (2003).** Exact arbitrage, well-diversified portfolios and asset pricing in large markets. *Journal of Economic Theory*, Vol. 110, No. 2, pp. 337–373. DOI: 10.1016/S0022-0531(03)00038-3.
11. **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.
12. **Ladrón de Guevara, R., Torra, S., & Monte, E. (2018).** 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, *Computación y Sistemas*, Vol. 22, No. 4, pp. 1049–1064. DOI: 10.13053/CyS-22-4-3083.
13. **Lagona, F. & Padovano, F. (2007).** A nonlinear principal component analysis of the relationship between budget rules and fiscal performance in the European Union. *Public Choice*, Vol. 130, No. 3-4, pp. 401–436. DOI: 10.1016/S0022-0531(03)00038-3.
14. **Lendasse, A., Lee, J., de Bodt, E., Wertz, V., & Verleysen, M. (2001).** Dimension reduction of technical indicators for prediction of financial time series—Application to the BEL20 Market Index. *European Journal of Economic and Social*

- Systems*, Vol. 15, No. 2, pp. 31–48. DOI: 10.1051/ejess:2001114.
15. **Marin, J. & Rubio, G. (2001).** *Economía Financiera*. Barcelona: Antoni Bosch.
 16. **Martin, E. & Morris, A. (1999).** Artificial neural networks and multivariate Statistics. *Statistics and Neural Networks Advances in the interface*, pp. 195–251.
 17. **Nieto, B. (2001).** *La valoración de activos en el mercado Español de valores: Tres ensayos*. Ph.D. Dissertation, University of Alicante. pp. 1–207.
 18. **Ravi, V. & Pramodh, C. (2008).** Threshold accepting trained principal component neural network and feature subset selection: Application to bankruptcy prediction in banks. *Applied Soft Computing*, Vol. 8, No. 4, pp. 1539–1548. DOI: 10.1016/j.asoc.2007.12.003.
 19. **Roll, R. & Ross, S. (1980).** An empirical investigation of the Arbitrage Pricing Theory. *Journal of Finance*, Vol. 35, No. 5, pp. 1073–1103. DOI: 10.1111/j.1540-6261.1980.tb02197.x.
 20. **Scholz, M. (2002).** *Nichtlineare Hauptkomponentenanalyse auf Basis neuronaler Netze*. Diplomarbeit, Berlin: Humboldt-Universität zu Berlin, Institut für Informatik.
 21. **Scholz, M. (2006).** *Approaches to analyse and interpret biological profile data*. Ph.D. Dissertation, Postdam University.
 22. **Scholz, M. (2006b).** *Nonlinear PCA toolbox for Matlab®*.
 23. **Scholz, M. (2007).** Analyzing periodic phenomena by circular PCA. S. Hochreiter & R. Wagner (eds.), *Proceedings of the Conference on Bioinformatics Research and Development*, pp. 38–47. DOI: 10.1007/978-3-540-71233-6_4.
 24. **Scholz, M. (2008).** *Non-linear PCA website*.
 25. **Scholz, M. & Vigarío, R. (2002).** Nonlinear PCA: a new hierarchical approach. M. Verleysen (ed.), *Proceedings of the European Symposium on Artificial Neural Networks*, pp. 439–444. Bruges: ESANN.
 26. **Scholz, M., Kaplan, F., Guy, C., Kopka, J., & Selbig, J. (2005).** Non-linear PCA: a missing data approach. *Bioinformatics*, Vol. 21, No. 20, pp. 3887–3895. DOI: 10.1093/bioinformatics/bti634.
 27. **Scholz, M., Fraunholz, M., & Selbig, J. (2007).** Nonlinear principal component analysis: Neural network models and applications. In **A. Gorban et al (eds.)**, *Principal manifolds for data visualization and dimension reduction*, pp. 44–67. DOI: 10.1007/978-3-540-73750-6_2.
 28. **Sun, X. & Ni, Y. (2006).** Recurrent neural network with kernel feature extraction for stock prices forecasting. *Proceedings of the International Conference in Computational Intelligence and Security*, Vol. 1, pp. 903–907. DOI: 10.1109/ICCIAS.2006.294269.
 29. **Tudón, J., Morales, R., Ramírez, R., Garza, L., & Vargas, A. (2011).** Fault detection in a Heat Exchanger, comparative Analysis between Dynamic Principal Component Analysis and Diagnostic Observers. *Computación y Sistemas*, Vol. 14, No. 3, pp. 269–282.
 30. **Weigang, L., Rodrigues, A., Lihua, S., & Yukuhiro, R. (2007).** Nonlinear Principal Component Analysis for Withdrawal from the Employment Time Guarantee Fund. *Computational Intelligence in Economics and Finance*, Vol. II, 75–92. DOI: 10.1007/978-3-540-72821-4_4.
 31. **Zangari, P. (2003).** Equity risk factor models. In **R. Litterman (ed.)**, *Modern Investment Management: An equilibrium approach*.
 32. **Zivot, E. & Wang, J. (2003).** *Modeling financial time series with S-Plus®*.

Article received on 21/10/2018; accepted on 06/02/2019.
Corresponding author is Rogelio Ladrón de Guevara Cortés.