Is the predictability of emerging and developed stock markets really exploitable?

David Moreno [a,*], Ignacio Olmeda [b]

[a] *Departamento de Economía de la Empresa, Universidad Carlos III, C/Madrid 126, 28903 Getafe Madrid, Spain*
[b] *Dpto. de Ciencias de la Computación, Universidad de Alcalá Madrid, Spain*

Abstract

A number of recent papers have analyzed the degree of predictability of stock markets. In this paper, we firstly study whether this predictability is really exploitable and secondly, if the economic significance of predictability is higher or lower in the emerging stock markets than in the developed ones. We use a variety of linear and nonlinear Artificial Neural Networks models and perform a computationally demanding forecasting experiment to assess the predictability of returns. Since we are interested in comparing the predictability in economic terms we also propose a modification in the nets' loss function for market trading purposes. In addition, we consider both explicit and implicit trading costs for emerging and developed stock markets. Our conclusions suggest that, in contrast to some previous studies, if we consider total trading costs both the emerging as well as the developed stock returns are clearly nonpredictable. Finally, we find that Artificial Neural Networks do not provide superior performance than the linear models.

Keywords: Finance; Forecasting; Emerging stock markets; Artificial neural networks

1. Introduction

During the last decade the equity developing markets in Latin America, Eastern Europe, Asia and Africa have achieved a great relevance, both from an academic point of view as well as from the point of view of a professional investor. This growing attention is due to the fact that returns from emerging markets are generally higher than returns from developed markets and display a low correlation among them, offering to investors new possibilities to enhance portfolio performance in

☆ We would like to thank to Paulina Marco, Natividad Blasco and participants in the IX Forum of Finances (2001, Spain). We also thank to Guillermo Llorente and participants in the Finance Workshop organized by Analistas Financieros Internacionales (2000) and seminars at Villanova University (2002, USA) and Universidad Carlos III (2003, Madrid) for their helpful com ments. All errors are the sole responsibility of the authors. Moreover, David Moreno acknowledges financial support from Ministerio de Ciencia y Tecnología grant SEJ2004 01688/ECON.

* Corresponding author. Tel.: +34 916245794; fax: +34 916249607.

E mail address: jdmoreno@emp.uc3m.es (D. Moreno).

terms of risk-adjusted-return (see Levy and Sarnat, 1970; Harvey, 1995; Solnik, 1995 or Kohers and Pandey, 1998, among others). The growing interest has been also motivated because these markets have experienced important liberalization processes along the late 1980s and early 1990s (see Bekaert et al., 2003 for a review) which has facilitated the access of foreign investors.

The academic literature has also reflected this growing interest: as it is well known, one of the most relevant topics in Finance has been the study of the predictability of stock market returns, and an immediate open question is whether emerging markets are more or less predictable than the developed ones. While there is a number of studies on the predictability of returns in developed markets (e.g. Campbell, 1987; Breen et al., 1990; Pesaran and Timmermann, 1994, 1995 for US markets and Clare et al., 1994; Fama and French, 1998; Pesaran and Timmermann, 2000 for some other markets,[1] or more recently Lewellen, 2004 or Guo, 2006 among others), for emerging markets, the number references has been much smaller and, equally, non-conclusive. For example, Harvey, 1995; Fama and French, 1998; Harvey et al., 2000 or van der Hart et al., 2003 have found evidence in favor of predictability while other studies have found empirical evidence against it (e.g. Urrutia, 1995 or Chang et al., 2004).[2]

In Table 1 we present a summary of just some recent studies analyzing the predictability of emerging stock markets.[3] As it can be seen, most of the studies suggest some degree of predictability. Note, however, that most of these studies do not take into account the effect of transaction costs even though, as some relevant research suggest (e.g. Bekaert et al., 1997; Harvey et al., 2000; Domowitz et al., 2001 or Chang et al., 2004) trading costs in these markets are of a such magnitude that the conclusions could be different it they would have been explicitly considered. In this paper, we will consider both explicit (the direct costs of trading, such as broker commission costs, taxes, etc.) and implicit trading costs (representing indirect trading costs, the major one being the price impact of the trade) for emerging and developed stock markets. It must be noted this is an important point given that in many markets, especially in the developing ones, the implicit costs are even higher than the explicit costs, thus not considering the total cost of trading could bias the results.

Our approach will try to reasonably mimic the problem faced by an American[4] institutional investor or mutual fund manager who wants to know what will be the next day or next week return of a particular market, and who employs the past history of returns in the market as well as of other reference markets (in particular, USA, Japan, United Kingdom and Germany). To do so, we will conduct forecasting experiments using a wide database, several information sets and model specifications, which will allow us to assess the robustness of our results. We will also adopt a linear and nonlinear perspective, the inclusion of nonlinear forecasts is deemed important as there is now considerable evidence of nonlinearity in stock market returns (e.g. Hsieh, 1991).

Even though there is a huge number of parametric and nonparametric techniques which could be used to model nonlinearity, in this study we employ one of the most powerful ones: Artificial Neural Networks (ANNs hereafter). ANNs have been used for a wide variety of issues and problems but forecasting is one of the main applications (see Zhang et al., 1998 for a review). There are some features of the ANNs that made it a proper tool for evaluating the predictability of stock returns. First, ANNs have the capability of generalization in the sense that after a learning period from a sub-sample they can be employed to forecast another unseen subsample, even if some noise is present in the data. Second, as some researchers have shown (e.g. Hornik et al., 1989; Hornik, 1991) ANNs can approximate any continuous function to an arbitrary level of accuracy so that they can be considered much more flexible than the traditional statistical methods. Moreover, ANNs are inherently nonlinear and nonparametric so that a specific functional form is not needed to be specified.

[1] Note, however, that some other studies have also pointed out the absence of this predictability (e.g. Nelly and Weller, 2000 or Ang and Bekaert, 2001).

[2] We are grateful to an anonymous referee for helping us to improve substantially the motivation of the paper.

[3] It must be noted that this table is by no means complete and some studies could be not included.

[4] We only consider indexes expressed in US dollars and not in each one of the local currencies of each of the countries analyzed. Although an analysis from the perspective of a home investor would also be interesting, it would be more difficult to compare the results with those of previous studies.

Table 1
Summary of studies on predictability in different developing stock markets

Researchers	Frequency data	Number of Stock market analyzed	Do they find any predictability?	Use of transaction cost
Achour et al. (1999)	Monthly, quarterly and semiannual (1988–1998)	1 (Malaysia)	Yes	No
Achour et al. (1998)	Monthly, quarterly and semiannual (1988–1998)	3	Yes	No
Barry et al. (2002)	Monthly (1985–2000)	35	Yes	No
Chang et al. (2004)	Daily (1991–2004)	11 emerging markets and 2 developed markets	Yes (without transaction cost) No (with transaction cost)	Yes
Fama and French (1998)	Monthly (1975–1995)	16 emerging markets and 13 developed markets	Yes	No
Harvey (1995)	Monthly (1976–1992)	20	Yes	No
Harvey et al. (2000)	Monthly (1992–1997)	9	Yes	No
Rouwenhorst (1999)	Monthly (1982–1997)	20	Yes	No
Urrutia (1995)	Monthly (1975–1991)	4	No	No
van der Hart et al. (2003)	Monthly (1985–1999)	32	Yes	Yes
Milobedzki (2004)	Daily (1995–1999)	1 (Warsaw)	Yes	No

Since we are interested in comparing the predictability in economic terms, it will be necessary to employ a loss function compatible with this objective. As White (1989) early pointed out, it is extremely important to employ the same loss function to build the net and to evaluate its forecasts. A common error in some studies (e.g. Harvey et al., 2000 or Jasic and Wood, 2004) is to employ a loss function (generally, the mean squared error) to build the net and then to evaluate its forecast using another different one. Obviously this way to proceed is inconsistent since the parameters obtained could not be optimal under the new loss function. For this reason, in this paper we will employ the same economic measure to train and to evaluate the net assuring consistency. Unfortunately, the modification of the loss function for market trading purposes extremely complicates the training process since the use of arbitrary performance measures do not allow to employ standard optimization algorithms such as *backpropagation*. To overcome this problem, in this paper we employ a heuristic optimization algorithm (the Genetic Algorithm, GA) to optimize the economic performance of the trained net.

Specifically, the objective of this paper is twofold: First, using a wide number of series we will try to obtain general conclusions about the comparative degree of predictability between emerging and developed markets returns; at the same time, we will also evaluate the usefulness of employing complicated nonlinear models (ANNs) against simpler autoregressive ones, and second, we will try to evaluate whether this predictability is really exploitable in economic terms from the point of view of an investor by considering the total transaction costs. Our contribution to the existing literature comes from the use of both linear and nonlinear models as well as the employment of a loss function which incorporates implicit and explicit transactions costs consistent with real economic performance of the models. Our paper also gives some evidence on the issue of the degree of predictability of developed and emerging stock markets.

The remainder of the paper is organized as follows. In the next section we will briefly describe the models employed as well as the method used to test the significance of the predictions. In this section we also give a brief introduction to the nonparametric model used (Artificial Neural Networks). The third section describes the period of study as well as the data used. In this section, we also detail the procedure employed and summarize the main empirical results. Some concluding remarks are offered in the final section.

2. Methodology employed

In this study we basically follow a quite simple but extremely computationally intensive approach: we build a variety of linear and nonlinear models estimated using several information sets and use them to obtain forecasts of a significant number of series and along an extensive time span. Subsequently we employ test procedures to detect the possible asymmetries between models, information sets as well as specifications.

To produce our forecasts we use four alternative specifications and several information sets. Our first specification is an autoregressive model of the form:

$$R_t^i = \alpha + \sum_{j=1}^{p_i} \beta_j^i R_{t-j}^i + \varepsilon_t^i, \tag{1}$$

where R_t^i is the return at time t of the stock market index of country i, p_i is the number of lags in the model and ε_t^i is a white noise process. In the second specification we employ bivariate autoregressive models:

$$R_t^i = \alpha + \sum_{j=1}^{p_i} \beta_j^i R_{t-j}^i + \sum_{k=1}^{q_i} \gamma_k^i R_{t-k}^s + \varepsilon_t^i, \tag{2}$$

where R_t^s is the return at time t of the stock market index of country s, different from i, p_i and q_i are the number of lags in the model and ε_t^i is a white noise process. Note that (1) is a nested model of (2). It must be noted that authors have been extremely careful with the determination of the proper lag returns that are included in each model, to avoid a problem of overlapping trading periods between countries in different time zones.[5]

Our last two specifications are nonlinear versions of (1) and (2). If we suppose that the returns are predictable using information of past returns from the markets, then we are implicitly assuming that:

$$R_t^i = f(R_{t-1}^i, R_{t-2}^i, \ldots, R_{t-p_i}^i, R_{t-1}^s R_{t-2}^s, \ldots, R_{t-q_i}^s, \Theta) + \varepsilon_t^i, \tag{3}$$

where Θ is a vector of parameters and f is the expectation of R_t^i conditional on $R_{t-1}^i, R_{t-2}^i, \ldots, R_{t-p_i}^i, R_{t-1}^s, R_{t-2}^s, \ldots, R_{t-q_i}^s$. Under the *parametric* point of view, f is a known function and the problem of estimating the parameter vector Θ arises. When f is unknown (as is usually the case), one needs to employ flexible functional forms to approximate this function. In this case, we say that we are adopting a *nonparametric* point of view.

There are many techniques that can be used to approximate f in a nonparametric manner (e.g. Splines, Nearest Neighbours, etc.). In this study we employ one of the most powerful, the Artificial Neural Networks (ANNs hereafter). ANNs have some properties that make them a convenient tool for forecasting (see Zhang et al., 1998 for a complete review), among them: (i) their capability of generalization; (ii) they can approximate any continuous function to any desired level of accuracy; (iii) ANNs are inherently nonlinear and nonparametric and, therefore, a specific functional form is not needed to be specified. In our specific case, the use of a nonparametric methods is particularly convenient: since a wide number of different equity markets are analyzed. It could be quite logical to think that different functional forms should be employed in each of the cases complicating the choice and comparability of the models and making the problem of misspecification more severe. Our choice of ANNs among other nonparametric methods is purely practical since any consistent nonparametric model would be able to approximate the underlying data generating process of the returns. However, it should be noted that the statistical properties of ANNs are much better known (e.g. Hornik et al., 1989) than for other nonparametric models; also, the extended use of ANNs in forecasting problems as well as the availability of thoroughly tested training algorithms motivate our choice. For reasons of brevity we will describe ANNs models succinctly and we refer the interested reader to the references given.[6]

An ANN is composed of a number of *processing units* which are hierarchically organized in *layers*. The input layer consists of a set of nodes that receive the information from the outside world.

[5] In the case of Australia and New Zealand we take into account that these markets are in the pre opening period in the t day when the American market is still open in its $t-1$ day. Thus, if we would include the information of USA lagged one period in the bivariate models, we would be incorporating information that is posterior to the opening of the local stock market.

[6] Kuan and White (1994) provide an excellent introduction to ANN. Also, White (1989) and Cheng and Titerington (1994) provide an introduction from a statistical perspective.

The *hidden layer* processes the information while the output layer sends the signal to the outside. The most widely used structure is that of a *feedforward neural net* in which the information is hierarchically processed in a single way from the input layer to the output through the hidden layer(s). Other structures allowing feedback are also possible, but we will not consider them here to maintain the paper relatively compact.

The units are connected through *synaptic weights* which determine quantitatively the influence of one unit on the other. A unit has an inhibitory or excitatory effect on the other depending on whether the sign of the corresponding weight is positive or negative. The set of interconnecting weights between units i and j, (W_{ij}), is known as the *weighting matrix*. Following Kuan and White (1994), the process of transforming inputs to outputs in a feedforward ANN with r inputs, one hidden layer of q units and a single output unit can be parameterized in the following way:

$$\hat{f}(x, W) = F\left(\beta_0 + \sum_{j \quad 1}^{q} \beta_j G(x'\gamma_j)\right), \tag{4}$$

where, $\hat{f}(x, W)$ is the *output* of the net, $x \quad (1, x_1, x_2, \ldots, x_r)'$ represents the input (the "1" corresponds to the *bias* in a traditional model), $\gamma_j = (\gamma_{j0}, \gamma_{j1}, \ldots, \gamma_{ji}, \ldots, \gamma_{jr})' \in \Re^{r+1}$ are the weights from the input to the hidden layer, β_j represents the weights from the hidden to the output and $F : \Re \to \Re$ and $G : \Re \to \Re$ are the *activation* functions of the output and hidden units, respectively (generally, the logistic function $G(a) \quad 1/(1 + \exp(-a))$). As we can see from the above expression, $G(x'\gamma_j)$ corresponds to the well-known *logit* model of binary response. Hornik et al. (1989) have shown that an ANN with a single hidden layer with enough hidden logistic units and linear outputs can approximate arbitrarily well any measurable function.

The problem is, obviously, how to find the weights that index the functions involved in the above expression. This procedure is usually referred to as *learning*. The learning of the network can be understood as a trial and error procedure that allows us to find the parameters that minimize the errors of the net, that is, given the desired output (Y_t), and the actual output of the net $(\hat{f}(x_t, W))$, the problem comes to minimize the error between the actual and desired output $(e_t = e(\hat{f}(x_t, W), Y_t))$ along a *training set* of examples.

Formally, the process of *learning* consists of solving a nonlinear least squares problem, for which many methods can be applied. The most widely used method is that of *backpropagation of errors* ("Backpropagation"), based on the stochastic approximation algorithm of Robbins and Monro (1951).[7]

The main problem in the implementation of an ANN model is their flexibility which can lead to *overtraining*, which refers to the problem of obtaining structures with low errors along the training set but with high errors along the *testing set*. This problem generally applies to nonparametric models and is caused by an excessive number of parameters of the model in relation to the complexity of the problem and sample size. Also, low parameterized networks would be unable to capture the functional relationship between input and output. There are several ways to remedy this problem but in this paper we will adopt a simple solution by constraining learning to an arbitrary bound depending on the error of an equivalent linear model.

3. Database and results

Our database consists of 49 MSCI (Morgan Stanley Capital International) indexes, expressed in US dollars, covering the period from March 1995 to March of 2001 (1560 daily observations). Twenty one of the indexes correspond to developed markets and the remaining 28 to emerging markets.[8]

---

[7] The method consists of an iterative procedure in which the new weights are obtained according to the following expression:

$$W_{t+1} \quad W_t + a\nabla\hat{f}(x_t, W_t)(Y_t \quad \hat{f}(x_t, W_t)),$$

where a is a constant named *learning rate* and $\nabla\hat{f}(x_t, W_t)$ is the gradient of the output at time t, $\hat{f}(x_t, W_t)$, with respect to the weights, W_t.

[8] The emerging markets analyzed are China (CHI), India (IND), Indonesia (INO), Korea (KOR), Malaysia (MAL), Pakistan (PAK), Philippines (PHI), Sri Lanka (SRI), Taiwan (TAW), Thailand (THA), Argentina (ARG), Brazil (BRA), Chile (CHE), Colombia (COL), Mexico (MEX), Peru (PER), Venezuela (VEN), Czech Republic (CZE), Egypt (EGY), Greece (GRE), Hungary (HUN), Israel (ISR), Jordan (JOR), Morocco (MOR), Poland (POL), Russia (RUS), South Africa (SOU) and Turkey (TUR). The developed markets are Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DEN), Finland (FIN), France (FRA), Germany (GER), Hong Kong (HON), Ireland (IRE), Italy (ITA), Japan (JAP), Luxembourg (LUX), Netherlands (NET), Norway (NOR), Portugal (POR), Singapore (SIP), Spain (SPA), Sweden (SWE), Switzerland (SWI), United Kingdom (UKG) and United States (USA).

Since this database is well known to the financial researchers we do not provide a full description here.[9] The returns of the markets are calculated as first differences of log prices. The main statistics of the raw data are shown in Table 2. As it is commonly found, we document evidence of leptokurtosis and asymmetry in the distribution of the returns, which allow the rejection of normality by means of the Jarque Bera test. We also find strong evidence of ARCH effects (the only exception is Jordan) by applying Engle's test of four lags.

To test whether linear structure is present in the data, we adopt a model selection perspective. We estimate autoregressive models with up to 10 lags and choose the model which minimizes the Akaike Information Criterion. Then, for the selected model, we check whether there is serial correlation in the data. If there is not serial correlation the model should not include any additional lagged terms. The results of this procedure are reflected in the first two lines (for each country) of Table 3. Note that with the exception of Luxembourg, all the markets show some low-order autocorrelation.

Having found some evidence of linear structure we proceed to test for neglected nonlinearities. To do so, we filter each of the series by the models selected in the first step. Since, in some cases, the estimated model is unable to remove all the linear structure in the data (the Ljung Box test rejects the null of no autocorrelation at 10% for the first ten lags), we employ an alternative autoregressive model where the lag order is the minimum so that the Ljung Box test is unable to reject the null of no autocorrelation in the residuals. The results are shown, for each of the countries, in the following two lines. Note that for five emerging markets (India, Thailand, Brazil, Czech Republic and South Africa) and for three developed markets (Italy, Spain and Sweden), the chosen model is unable to remove all the linear structure, therefore, we employ the alternative model.

After filtering through the chosen model, we calculate the BDS statistic (Brock et al., 1996) for the residuals. Following the suggestions by Hsieh (1991), the proximity parameter is set equal to the standard deviation of the series[10] while the embedding dimension, m, goes from 2 to 5. As we can see in Table 4, in all the cases there is a clear rejection of the null, which can be interpreted as evidence in favor of neglected nonlinearity or nonstationarity of the series (the only exception is Israel for m 2). As the time span is relatively short we interpret this rejection as evidence in favor of nonlinearity (in mean or variance).

As a conclusion of this preliminary analysis, we can say that there exists statistical evidence in favor of linear and nonlinear structure in each of the daily and weekly series.[11] Obviously, one cannot conclude that this evidence should be exploitable to forecast future returns. For example, nonlinearity could be in variance motivating the rejection of the null by the BDS test. If this were the case, nonlinear in mean models could be useless. This motivates the forecasting experiments in the next sections.

3.1. Linear forecasts

To compare the linear univariate predictability of the markets for daily observations, we use the following procedure. For each one of the markets and at any moment of time we consider the last 250 observations, and then we estimate autoregressive linear models (1) with a maximum length lag of ten and select the model which minimizes the Akaike Information Criterion (AIC). This model is used to forecast the next observation and the difference against the observed return is computed. After doing this we roll the window one day ahead, keeping the size of the window constant, and then proceed as before until the end of the sample is reached. For the bivariate models (2) we also considered, respectively, lagged returns from the United States, Japan, Germany, and the United Kingdom.[12]

[9] It is important to note that every index represents at least 60% of the capitalization of the whole national stock market. In addition, we must also note the convenience of employing MSCI database in studies where there exist comparisons between emerging and developed stock markets, since in the MSCI database both emerging and developed indexes are computed according to the same criterions.

[10] The BDS statistic was also computed for a proximity parameter between 0.5 and 2 times the standard deviation of the series but the results were the same as those reported here.

[11] Although we have not shown the BDS results for weekly data, they are very similar.

[12] We recognize, however, that other available variables could influence the evolution of future returns in this time frequency, for example short or long term interest rates and exchange rates, but to keep the study reasonably compact, we will not consider them here, leaving their study for future research. In some sense, this piece of work can be interpreted as an evaluation of market efficiency in its weakest form.

Table 2
Descriptive statistics of the daily series

	Mean (%)	Std. Dev.	Max. (%)	Min. (%)	Skewness	Kurtosis	Prob. J B	Arch test
CHI	0.079	0.023	0.076	0.061	0.199	7.697	0.00	0.000
IND	0.015	0.018	0.055	0.043	0.077	5.225	0.00	0.000
INO	0.119	0.038	0.137	0.111	0.975	22.563	0.00	0.000
KOR	0.051	0.031	0.029	0.030	0.429	12.391	0.00	0.000
MAL	0.051	0.027	0.115	0.056	0.638	41.667	0.00	0.000
PAK	0.075	0.022	0.057	0.087	0.533	10.174	0.00	0.000
PHI	0.078	0.020	0.034	0.046	1.387	20.001	0.00	0.000
SRI	0.090	0.013	0.047	0.031	0.013	7.600	0.00	0.000
TAW	0.018	0.019	0.058	0.067	0.054	6.083	0.00	0.000
THA	0.141	0.027	0.055	0.050	0.621	8.018	0.00	0.000
ARG	0.016	0.021	0.047	0.081	0.029	9.876	0.00	0.000
BRA	0.022	0.023	0.048	0.044	0.010	10.502	0.00	0.000
CHE	0.027	0.012	0.028	0.030	0.092	6.928	0.00	0.000
COL	0.073	0.014	0.053	0.048	0.346	11.862	0.00	0.000
MEX	0.050	0.021	0.063	0.047	0.358	14.641	0.00	0.000
PER	0.002	0.015	0.044	0.026	0.326	10.152	0.00	0.000
VEN	0.012	0.029	0.067	0.046	8.956	23.30	0.00	0.000
CZE	0.009	0.015	0.035	0.029	0.117	5.188	0.00	0.000
EGY	0.000	0.014	0.049	0.036	0.434	7.291	0.00	0.000
GRE	0.038	0.020	0.061	0.067	0.063	5.747	0.00	0.000
HUN	0.052	0.022	0.054	0.033	0.573	12.340	0.00	0.000
ISR	0.036	0.017	0.027	0.033	0.495	8.218	0.00	0.000
JOR	0.029	0.008	0.029	0.030	1.406	16.060	0.00	0.112
MOR	0.023	0.007	0.023	0.012	0.567	10.117	0.00	0.000
POL	0.015	0.021	0.039	0.0423	0.208	5.525	0.00	0.000
RUS	0.058	0.040	0.097	0.083	0.289	9.180	0.00	0.000
SOU	0.029	0.015	0.037	0.039	0.769	10.609	0.00	0.000
TUR	0.001	0.034	0.088	0.097	0.210	10.307	0.00	0.000
AUT	0.017	0.011	0.020	0.026	0.377	5.616	0.00	0.000
BEL	0.020	0.011	0.023	0.026	0.208	6.674	0.00	0.000
CAN	0.039	0.013	0.026	0.020	0.788	9.266	0.00	0.000
DEN	0.045	0.012	0.024	0.027	0.317	4.764	0.00	0.000
FIN	0.095	0.025	0.060	0.046	0.365	8.975	0.00	0.000
FRA	0.040	0.012	0.025	0.018	0.136	4.482	0.00	0.000
GER	0.032	0.013	0.022	0.026	0.201	5.145	0.00	0.000
HON	0.010	0.019	0.045	0.029	0.188	11.859	0.00	0.000
IRE	0.026	0.012	0.031	0.026	0.168	6.717	0.00	0.000
ITA	0.045	0.014	0.020	0.022	0.031	4.691	0.00	0.000
JAP	0.022	0.015	0.035	0.031	0.523	7.366	0.00	0.000
LUX	0.083	0.012	0.024	0.032	1.796	16.535	0.00	0.000
NET	0.036	0.012	0.028	0.019	0.089	5.425	0.00	0.000
NOR	0.010	0.013	0.027	0.030	0.388	7.987	0.00	0.000
POR	0.030	0.012	0.018	0.018	0.131	6.055	0.00	0.000
SIP	0.026	0.016	0.046	0.037	0.321	10.502	0.00	0.000
SPA	0.062	0.013	0.019	0.022	0.218	5.983	0.00	0.000
SWE	0.049	0.017	0.053	0.036	0.179	7.142	0.00	0.000
SWI	0.038	0.011	0.023	0.022	0.149	6.223	0.00	0.000
UKG	0.028	0.010	0.022	0.24	0.123	4.400	0.00	0.000
USA	0.054	0.011	0.034	0.029	0.373	7.044	0.00	0.000

The table shows the descriptive statistics of each of the countries, for daily data. They are mean (%), standard deviation (std. dev.), maximum (max.), minimum (min.), skewness (the third central moment divided by the cube of the standard deviation), kurtosis (measured as the fourth central moment divided by the square of the variance of the data), Jaque Bera (a statistical test for testing whether the series is normally distributed), its *p* value is shown. The last column is the *p* value for the Engle's autoregressive conditional heteroskedasticity (ARCH) test.

After obtaining 250 forecasts, for each of the specifications, we compute the two-tailed Diebold Mariano (DM) statistic (Diebold and Mariano, 1995) in order to test the difference in performance

Table 3
Test for autocorrelation and selection of linear models

	CHI	IND	INO	KOR	MAL	PAK	PHI	SRI	TAW	THA
Panel A. Emerging markets										
Lag	2	**1**	6	9	6	1	1	8	1	**1**
p Value	0.316	**0.072**	0.886	0.970	0.999	0.241	0.490	0.999	0.665	**0.081**
Alternative	None	4	None	None	None	None	None	None	None	2
p Value		0.117								0.116
	ARG	BRA	CHE	COL	MEX	PER	VEN	CZE	EGY	GRE
Lag	7	**2**	3	3	2	7	6	**1**	2	1
p Value	0.762	**0.043**	0.375	0.972	0.119	0.998	0.944	**0.029**	0.123	0.506
alternative	None	6	None	None	None	None	None	5	None	None
p Value		0.159						0.204		
	HUN	ISR	JOR	MOR	POL	RUS	SOU	TUR		
Lag	3	1	1	6	10	1	**1**	1		
p Value	0.301	0.854	0.487	0.986	1.000	0.231	**0.090**	0.152		
Alternative	None	None	None	None	None	None	3	None		
p Value							0.167			
	AUT	BEL	CAN	DEN	FIN	FRA	GER	HON	IRE	ITA
Panel B. Developed markets										
Lag	5	1	7	2	2	8	7	4	1	**1**
p Value	1	0.673	0.845	0.563	0.275	1	1	0.201	0.572	**0.089**
Alternative	None	None	None	None	None	None	None	None	None	3
p Value										0.34
	JAP	LUX	NET	NOR	POR	SIP	SPA	SWE	SWI	UKG
Lag	6	0	3	4	1	1	**2**	**1**	1	3
p Value	0.699	0.117	0.120	0.997	0.724	0.754	**0.063**	**0.059**	0.316	0.132
Alternative	None	None	None	None	None	None	3	5	None	None
p Value							0.116	0.354		
	USA									
Lag	3									
p Value	0.228									
Alternative	None									
p Value										

The table shows, for each of the countries, the model which minimizes the Akaike Information Criterion (AIC), as well as the final model chosen to filter the series. In the first *lag* line we show the number of lags of the model which minimizes the AIC, in the next line, *p value*, we show the significant level of the Ljung Box (L B) statistic computed on the residuals of the model (10 lags). If the null of no autocorrelation is rejected at the 10% level, the values appear boldface. In this case, we employ an alternative autoregressive model where the lag order is the minimum so that the L B test is unable to reject the null of no autocorrelation in the residuals. The lag length of this model as well as the *p* value of the L B tests appear in the following lines (*alternative* and *p value*).

against a random walk with a drift and finally, we take the means along all the rolling windows. Since this statistic has a standardized normal asymptotic distribution, a negative number smaller than -1.96 indicates better performance of the model against the random walk while a positive number bigger than 1.96 means the opposite. Implemented in this way, our procedure can be interpreted as evidence in favor or against a model that it is used by an investor who observed what happened last year, built his model using this information and evaluated it during one whole year. To check the robustness of the results we repeat the whole procedure, but employing another two different information sets that consist on rolling windows of 100 and 500 observations. An identical procedure is employed using weekly data but using, in this case, 50, 100 and 150 observations. The results are summarized in Table 5.

In the case of daily data, we can observe that univariate models are unable to beat the random walk in all cases. This is also the case when we consider

Table 4
BDS test for residuals of linear models

Countries	$m=2$	$m=3$	$m=4$	$m=5$
Panel A. Emerging markets				
ARG	8.64	9.97	10.79	11.50
BRA	12.62	14.69	17.11	19.69
CHE	8.94	10.94	12.38	13.70
CHI	11.73	15.31	18.34	21.76
COL	12.04	14.89	16.45	17.85
CZE	8.39	11.72	13.86	16.55
EGY	5.86	10.21	12.94	14.89
GRE	8.52	11.76	14.3	16.63
HUN	12.28	14.43	16.31	17.55
IND	5.47	6.64	7.53	8.31
INO	14.84	18.59	21.55	24.87
ISR	**1.64**	3.08	3.94	3.69
JOR	4.92	3.30	2.30	2.09
KOR	8.72	12.3	15.72	18.46
MAL	12.64	17.01	20.43	24.12
MEX	8.80	9.84	11.22	12.29
MOR	4.13	6.04	6.55	6.73
PAK	7.37	8.51	9.00	9.12
PER	7.15	8.69	9.96	10.85
PHI	12.00	15.97	19.19	22.3
POL	10.35	11.83	13.5	14.67
RUS	11.83	14.72	17.2	19.47
SOU	10.97	14.56	17.19	20.15
SRI	9.24	12.31	14.44	16.23
TAW	3.59	5.32	6.50	7.35
THA	11.35	14.29	17.19	19.59
TUR	5.64	7.16	8.36	9.33
VEN	9.70	12.64	14.79	17.46
Panel B. Developed markets				
AUT	5.64	7.56	8.53	9.63
BEL	6.97	8.89	10.76	12.41
CAN	8.35	11.42	13.55	16.13
DEN	6.27	8.03	9.79	11.82
FIN	7.08	9.22	10.33	11.84
FRA	3.74	4.26	4.71	5.08
GER	4.70	7.27	9.25	10.66
HON	8.70	10.24	12.08	14.00
IRE	7.10	8.53	9.84	10.71
ITA	3.59	5.39	6.70	7.52
JAP	3.00	5.21	6.87	8.60
LUX	6.62	7.26	7.91	8.42
NET	6.96	9.25	11.16	13.21
NOR	6.85	8.92	9.93	10.3
POR	8.14	10.29	12.19	14.4
SIP	11.43	15.97	19.65	23.04
SPA	5.48	6.93	8.51	9.37
SWE	5.20	6.24	7.14	7.81
SWI	5.31	7.11	8.68	9.75
UKG	3.95	5.99	7.65	8.68
USA	3.26	5.90	7.35	8.99

The table shows, for each one of the countries, the BDS statistic for testing the null that the residuals of linear models are iid (identical independently distributed). Under the null hypothesis the statistic, which is computed for a proximity parameter equal to the standard deviation of each of the series and embedding dimensions from 2 to 5, is distributed as a $N(0,1)$. Bold numbers indicate a failure to reject the null at the 10% level.

bivariate models, with the exception of the models which include lagged returns from the USA market. When the USA market is considered, we are able to beat the random walk for Germany in all the specifications and also for Poland and Finland (two specifications) and the Philippines, Ireland and the United Kingdom (last specification). Note that most of the countries for which there is evidence in favor of predictability, are developed countries. In the case of weekly data we are also unable to beat the random walk for all the countries and specifications. Generally, the values of the DM statistic for weekly data were higher than the ones for daily data (as it can be concluded from the maximum, minimum and mean values shown in the table) which essentially means that the linear models employed at a higher frequency seem to be more useful.

It must be noted that for daily data, the DM values are lower for bivariate models which include information from the USA market, but overall, we found that this effect is more important for developed markets. This finding agrees with the results published by Ang and Bekaert (2001). These authors point found that the US instruments are strong predictors of foreign equity returns in developed countries.[13] Moreover, our findings reinforces the empirical evidence found by some previous studies (e.g. Harvey, 1995 or Rouwenhorst, 1999) on the fact that global risk factors are unable to explain the returns of the emerging markets and local factors are much more relevant.[14]

Summarizing, we can conclude the following. Firstly, emerging and developed markets are generally unforecastable but developed markets seem to be easier. Secondly, linear models have little value if they do not include lagged information from the USA market. Thirdly, and unexpectedly, predictability increases with frequency, even though noise at a daily frequency is plausibly more important.[15]

[13] Note, however, that these authors employ completely different approaches.

[14] Indeed, for the period studied in this paper, emerging markets display a low correlation with developed markets (the correlation between developed markets is approximately 0.28 while between emerging and developed stock markets it is much lower, around 0.15).

[15] Another possible explanation for this fact is that for weekly specifications the power of the test could be significantly lower due to the reduction of the sample size.

Table 5
DM statistic for linear models

Size		Univariate	USA	JAP	GER	UKG
Daily						
100	Max.	1.691	2.250	2.865	2.427	2.789
	Min.	0.652	2.147	0.499	0.718	0.742
	Mean	1.010	0.169	1.342	1.219	1.231
	Countries	None	GER(d)	None	None	None
250	Max.	2.390	1.473	2.242	1.787	1.722
	Min.	1.121	2.579	1.129	1.128	1.246
	Mean	0.595	0.658	0.598	0.573	0.412
	Countries	None	POL(e) FIN(d) GER(d)	None	None	None
500	Max.	2.000	1.162	2.056	1.654	1.238
	Min.	1.381	3.583	1.715	1.701	1.849
	Mean	0.306	0.992	0.189	0.097	0.091
	Countries	None	POL(e) PHI(e) IRE(d) GER(d) FIN(d) UKG(d)	None	None	None
Weekly						
50	Max.	2.811	2.382	1.951	2.492	2.101
	Min.	0.265	0.308	0.323	0.375	0.360
	Mean	0.946	1.090	1.096	1.249	1.092
	Countries	None	None	None	None	None
100	Max.	1.653	1.951	2.101	2.022	2.014
	Min.	0.000	0.287	0.239	0.165	0.206
	Mean	0.702	1.100	1.056	1.031	0.994
	Countries	None	None	None	None	None
150	Max.	2.127	2.244	2.290	1.811	2.729
	Min.	0.000	0.000	0.000	0.003	0.000
	Mean	0.621	0.874	0.869	0.753	0.847
	Countries	None	None	None	None	None

In the table we show the maximum, minimum and mean values of the DM statistic for testing the null of equal accuracy against a drifted random walk for daily and weekly data and several window's sizes. In the line *countries* we show the developed (d) and emerging (e) markets for which the DM test rejects the null, that is, markets for which evidence of predictability is found.

3.2. Nonlinear forecasts

In this case, we employ ANNs to forecast next day or next week return. To reduce the computational burden we only consider five lags for the univariate models and three lags for each one of the regressors in the bivariate models. To make the results of this section comparable, we consider the same information sets as before (rolling windows of size 100, 250 and 500 for daily data; and lagged returns of the country in the univariate case and lagged returns of USA, Japan, UK and Germany in the bivariate case). The number of hidden units is 4 for rolling windows of size 100 and 250 and 8 for the other case. For weekly data the number of hidden units is 4 for windows of size 50 and 100 and 5 for windows of size 150.

The learning phase is implemented as follows[16]: for each one of the countries we generate 5000 initial ANN configurations and choose the one that has the minimum sum of squared residuals along the training set. Then, we train this net along

[16] We found that this approach allows us to obtain error rates along the training set between 85% and 95% of the error rate of the linear model, which are acceptable in the sense that they are low enough to be sure that the net has effectively learnt the training patterns but sufficiently high to prevent overtraining.

10.000 *epochs*[17] and compute the ratio of the sum of squared residuals (ssr) of the net against the ssr of a linear model (which is estimated with the same information set). If the net is, at least, 2% better than the linear model we accept this configuration and begin the learning phase, otherwise we repeat the procedure. After having found an acceptable initial configuration, we begin a dynamic process of estimation and prediction in the same manner as with the linear models: at each time step one observation is added and the first one in the training set is deleted. We then retrain the net for 100 epochs so that the information contained in this observation is "assimilated" by the net and finally, we compute the next day forecast. This procedure continues until the end of the sample is reached.

Since we are using nonparametric estimators in this experiment, the DM test cannot be directly applied because the distribution of the statistic is, in this case, unknown. Instead, we compute the ssr along the last 250 observations and compare it with the ssr of a random walk without drift. Assuming equal forecasting accuracy of the random walk, this ratio should be equal to one. A ratio smaller than one gives evidence in favor of predictability, while a ratio higher than one means that the random walk model dominates.

It could be argued that directional forecasts are indeed more relevant than mean squared error forecasts, since the decisions are usually taken based on the forecasted direction of the market, for this reason, to evaluate the performance of our predictions in a different way, we compute the proportion of correct forecasted directions; under the null of no predictability, this proportion should be equal to 0.5 (similar analysis is realized by Harvey et al., 2000). Finally, we employ a third measure which compares the performance of the models against a naive model which predicts that the next day direction of the market will be the same as the last one observed (which is essentially the model that a short memory agent with extrapolative expectations would use). To do this we compute the ratio of the mean proportion of directions correctly predicted by the linear and ANN models against the proportion obtained by using a naive model. Again, a number higher than one would denote better behavior than the naive model. The results for each of these measures are shown in Tables 6 8. Note that we also include the results obtained with the linear model which will allow us to compare the possible effects of neglected nonlinearity on the forecasts, as well as allowing us to compare the results obtained with this ratio with the ones of the DM statistic.

In terms of the ssr (Table 6), we are generally able to beat the random walk when we employ bivariate linear models which include information from the USA. Note also that the forecasts of nonlinear models are generally worse than the linear ones and that, when the market can be forecasted, the improvement in predictability is surprisingly greater for developed markets. When we employ weekly observations, the improvements over the random walk disappear: in none of the cases we are able to obtain useful forecasts. For the second performance measure (Table 7), we find that for almost all the specifications it is possible to beat the random walk. We do not find either substantial difference between daily and weekly data or between linear and nonlinear models. Finally, for the third measure (Table 8), we also find a substantial number of specifications for which it is possible to beat the random walk. Again this is especially true for developed markets. It is also apparent that it is more difficult for developed markets to be forecasted at a higher frequency and also that nonlinear models are slightly worse than the linear ones. Finally, note that for weekly data the improvement obtained by using information from the USA market becomes less relevant.

Overall we arrive to similar conclusions as the ones obtained before: firstly, incorporating past information from the USA market into the models is essential for forecasting purposes for higher frequency data. Secondly, nonlinear models are slightly worse than the linear ones. Thirdly, when we compare the predictability of the developing and developed stock markets, we find that developed markets are relatively more forecastable than the emerging ones when world information is used. These results are interesting and again suggest that local information is much more important in emerging markets.

3.3. Profitability comparison when transaction costs are introduced

As we can see from a close inspection of Table 1, the studies on predictability in stock markets generally have not taken into account the effect of

[17] One *epoch* is one learning cycle along all the observations in the training set.

Table 6
Mean ssr ratio in linear and neural network models

	Panel A. Emerging markets										Panel B. Developed markets									
	Univariate		Bivariate								Univariate		Bivariate							
			USA		JAP		GER		UKG				USA		JAP		GER		UKG	
	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET
Daily observations																				
100 obs, 4 hidden.	1.051	1.074	1.024	1.039	1.059	1.076	1.060	1.065	1.052	1.064	1.066	1.058	**0.983**	**0.997**	1.077	1.055	1.075	1.057	1.071	1.047
250 obs, 4 hidden.	1.005	1.072	**0.972**	1.000	1.009	1.075	1.007	1.075	1.002	1.065	1.022	1.040	**0.939**	**0.962**	1.023	1.038	1.024	1.045	1.017	1.038
500 obs, 8 hidden.	**0.996**	1.172	**0.959**	1.109	**0.996**	1.172	**0.992**	1.176	**0.986**	1.168	1.009	1.063	**0.935**	1.009	1.009	1.073	1.008	1.184	1.003	1.075
Mean	*1.017*	*1.106*	***0.985***	*1.049*	*1.021*	*1.108*	*1.020*	*1.105*	*1.013*	*1.099*	*1.032*	*1.054*	***0.952***	***0.989***	*1.036*	*1.055*	*1.036*	*1.096*	*1.030*	*1.053*
Weekly observations																				
50 obs, 3 hidden.	1.075	1.117	1.135	1.205	1.130	1.254	1.137	1.215	1.132	1.221	1.041	1.058	1.090	1.112	1.099	1.104	1.101	1.093	1.087	1.091
100 obs, 4 hidden.	1.066	1.250	1.126	1.527	1.111	1.599	1.127	1.579	1.118	1.509	1.055	1.110	1.120	1.239	1.105	1.262	1.108	1.265	1.103	1.217
150 obs, 5 hidden.	1.039	1.334	1.082	1.616	1.080	1.777	1.087	1.902	1.082	1.710	1.043	1.157	1.090	1.366	1.087	1.424	1.060	1.447	1.083	1.343
Mean	*1.060*	*1.234*	*1.114*	*1.449*	*1.107*	*1.543*	*1.117*	*1.565*	*1.111*	*1.480*	*1.046*	*1.108*	*1.100*	*1.239*	*1.097*	*1.263*	*1.090*	*1.268*	*1.091*	*1.217*

Mean ratio of the ssr (sum of squared residuals) of the linear and neural network models against a random walk (no drift). Bold numbers indicates that the corresponding model outperforms the random walk model and it happens when the mean ssr ratio is smaller than one. Underlined bold values denote that the net obtained a lower ssr ratio than both a linear model and the random walk.

Table 7
Proportion of correct directions using linear and neural network models

	Panel A. Emerging markets										Panel B. Developed markets									
	Univariate		Bivariate								Univariate		Bivariate							
			USA		JAP		GER		UKG				USA		JAP		GER		UKG	
	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET
Daily observations																				
100 obs, 4 hidden.	**0.518**	**0.519**	**0.535**	**0.528**	**0.518**	**0.518**	**0.518**	**0.520**	**0.519**	**0.519**	**0.514**	**0.515**	**0.558**	**0.554**	**0.509**	**0.504**	**0.513**	**0.511**	**0.512**	**0.517**
250 obs, 4 hidden.	**0.525**	**0.523**	**0.547**	**0.522**	**0.524**	**0.522**	**0.526**	**0.522**	**0.527**	**0.527**	**0.510**	**0.509**	**0.571**	**0.568**	**0.522**	**0.521**	**0.519**	**0.514**	**0.525**	**0.526**
500 obs, 8 hidden.	**0.529**	**0.521**	**0.551**	**0.548**	**0.527**	**0.520**	**0.530**	**0.519**	**0.535**	**0.524**	**0.514**	**0.510**	**0.571**	**0.564**	**0.524**	**0.515**	**0.521**	**0.509**	**0.531**	**0.521**
Mean	***0.524***	***0.521***	***0.544***	***0.533***	***0.523***	***0.520***	***0.525***	***0.520***	***0.527***	***0.523***	***0.511***	***0.511***	***0.569***	***0.565***	***0.518***	***0.513***	***0.518***	***0.512***	***0.525***	***0.523***
Weekly observations																				
50 obs, 3 hidden.	**0.517**	**0.511**	**0.522**	**0.516**	**0.523**	**0.530**	**0.523**	**0.519**	**0.514**	**0.524**	**0.535**	**0.535**	**0.534**	**0.521**	**0.522**	**0.538**	**0.525**	**0.543**	**0.532**	**0.533**
100 obs, 4 hidden.	**0.528**	**0.537**	**0.514**	**0.513**	**0.530**	**0.505**	**0.526**	**0.506**	**0.515**	**0.507**	**0.515**	**0.521**	**0.517**	**0.516**	**0.501**	**0.508**	**0.528**	**0.510**	**0.520**	**0.509**
150 obs, 5 hidden.	**0.532**	**0.528**	**0.518**	**0.506**	**0.531**	**0.505**	**0.533**	**0.505**	**0.515**	**0.526**	**0.507**	0.489	**0.501**	0.495	0.499	0.486	**0.525**	0.494	**0.502**	**0.516**
Mean	***0.526***	***0.525***	***0.518***	***0.512***	***0.528***	***0.513***	***0.527***	***0.510***	***0.515***	***0.519***	***0.519***	***0.515***	***0.517***	***0.511***	***0.507***	***0.511***	***0.526***	***0.516***	***0.518***	***0.519***

Proportion of correct forecasted directions of the linear and neural network models. A number higher than 0.5 indicates that the corresponding model outperforms the random walk model and it is marked by boldface numbers.

Table 8
Mean ratio of correct directions against a naive model

	Panel A. Emerging markets										Panel B. Developed markets									
	Univariate		Bivariate								Univariate		Bivariate							
			USA		JAP		GER		UKG				USA		JAP		GER		UKG	
	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET	Linear	NET
Daily observations																				
100 obs, 4 hidden.	0.996	0.997	**1.026**	**1.012**	0.995	0.996	0.993	0.997	0.996	0.996	0.995	0.997	**1.083**	**1.074**	0.987	0.978	0.993	0.989	0.994	**1.003**
250 obs, 4 hidden.	**1.003**	**1.001**	**1.045**	0.996	**1.003**	0.999	**1.005**	0.997	**1.006**	**1.007**	0.985	0.983	**1.100**	**1.095**	**1.008**	**1.005**	0.998	0.989	**1.013**	**1.014**
500 obs, 8 hidden.	**1.005**	0.990	**1.047**	**1.040**	1.000	0.989	**1.006**	0.986	**1.017**	0.996	0.991	0.985	**1.099**	**1.085**	**1.012**	0.994	**1.003**	0.980	**1.023**	**1.003**
Mean	***1.001***	*0.996*	***1.039***	***1.016***	*0.999*	*0.995*	***1.001***	*0.993*	***1.006***	*1.000*	*0.990*	*0.988*	***1.094***	***1.085***	***1.002***	*0.992*	*0.998*	*0.986*	***1.010***	***1.007***
Weekly observations																				
50 obs, 3 hidden.	0.968	0.954	0.977	0.964	0.978	0.990	0.979	0.972	0.962	0.979	**1.066**	**1.063**	**1.063**	**1.035**	**1.038**	**1.069**	**1.048**	**1.081**	**1.054**	**1.058**
100 obs, 4 hidden.	0.982	0.998	0.957	0.954	0.986	0.939	0.979	0.942	0.960	0.941	**1.023**	**1.033**	**1.025**	**1.025**	0.993	**1.006**	**1.052**	**1.014**	**1.027**	**1.008**
150 obs, 5 hidden.	**1.004**	0.989	0.978	0.953	0.998	0.952	**1.005**	0.952	0.973	0.989	**1.068**	**1.028**	**1.062**	**1.046**	**1.055**	**1.021**	**1.112**	**1.039**	**1.055**	**1.078**
Mean	*0.985*	*0.980*	*0.971*	*0.957*	*0.987*	*0.960*	*0.988*	*0.955*	*0.965*	*0.970*	***1.052***	***1.041***	***1.050***	***1.035***	***1.029***	***1.032***	***1.071***	***1.045***	***1.045***	***1.048***

Mean ratio of the number of correct forecasted direction of the linear and neural network models against a naive model which uses the last direction. A number higher than one show that the corresponding model outperforms the naive model and it is indicated in bold numbers. Underlined bold values denote that the net's forecast outperforms to both a linear model and the naive model.

transaction costs even though some researchers (e.g. Bekaert et al., 1997; Harvey et al., 2000; Domowitz et al., 2001 or Chang et al., 2004) have correctly indicated that the magnitude of costs in many of the markets would make trading unprofitable. In this section, we introduce the total costs of trading for each of the markets in the investments strategies by explicitly including the costs of execution.[18]

Following Domowitz et al. (2001) we consider both explicit and implicit trading cost. The explicit costs are the direct costs of trading, such as broker commission costs, taxes, etc. The implicit costs represent indirect trading costs, being the most important the price impact of the trade (which can be important in some emerging markets where some illiquid stocks could be traded).

In this set of experiments, we proceed similarly as before: we build ANN models and employ them to forecast the next return. We then shift the window and repeat the experiment until the end of the sample is reached. The innovation is that now we will optimize (train) the nets using a measure of profit instead of the standard mean squared error overcoming, as we mentioned, a common error in other empirical studies (e.g. Harvey et al., 2000 or Jasic and Wood, 2004).

We can summarize our procedure as follows: let us consider observations $r_1, r_2, \ldots, r_{\text{train}}$ as the training set of the net, and let $f_1, f_2, \ldots, f_{\text{train}}$ be the in-sample forecasts obtained with the net. As we have mentioned, in the preceding section we employed the mean squared error $\sum_i (r_i - f_i)^2$ as the loss function to be minimized in the training phase. Let us now assume that forecasts $f_1, f_2, \ldots, f_{\text{train}}$ produce a trading position whenever the expected return in absolute terms is bigger than the roundtrip cost, which would compensate at least the cost of trading (we do not consider opportunity costs) and let $rf_1, rf_2, \ldots, rf_{250}$ be the realized returns of the trading positions, that is, the actual return, when the position is taken, minus the roundtrip costs. We consider explicit and implicit trading costs, again, a common mistake in earlier studies is to take into account only the explicit ones. However, in many stock markets (especially emerging ones) the implicit cost are even higher, so not considering the whole trading costs could mean obtaining erroneous conclusions. A clear candidate of loss function is the geometric product of realized returns, $\prod_i (1 + rf_i)$. Note that, since we assume that short selling is permitted, rf_i can be positive (the directional forecast of the net is correct, therefore the return of the trade would be the one experienced by the market, r_i, in absolute terms), negative (the directional forecast of the net is wrong and the return of the trade would be minus the one of the market, r_i, in absolute terms) or zero (no trade occurs).

Obviously, the objective of the training phase is now to maximize the geometric return along the training set. Unfortunately, though justified, the modification of the loss function extremely complicates the practical implementation of the training process making it impossible to employ standard optimization algorithms (such as the *backpropagation* algorithm mentioned before) and forcing us to implement a heuristic optimization algorithm known as the Genetic Algorithm, GA to maximize the geometric return. A detailed description of GAs is beyond the scope of the paper and we refer to the excellent introduction of Vose (1999) to the interested reader.[19]

To avoid problems associated with overtraining, we proceed as follows. For each rolling window we train the net with the GA so as to maximize the geometric return. For each iteration of the algorithm we also compute the geometric return of the net for another out-of-sample validation set and, after learning has finished, we choose the net which has the highest geometric return along this validation set. Finally, this "optimal" net is used to forecast another different test set and the window is displaced one period in the future.[20] This procedure tries to mimic the position of a trader which has several forecasting alternatives which he has optimized using past data. Since he is interested in future and not past performance, he would compare these

[18] Given that Domowitz et al., 2001 does not provide these costs for all the stock markets considered here, we have to restrict our subsequent analyses to the markets where they are available Another possibility previously considered was to estimate these costs by taking the mean total cost of the markets of the same geographical area. Since these authors document a high variabil ity among the execution costs in countries belonging to the same region, we have discarded this possibility.

[19] The description of the method employed is relatively cum bersome so that we do not provide the technical details which can be obtained from the authors.

[20] Another final comment must be made, since both the nets as well as GAs depend on several parameters chosen by the experimenter, data snooping problems are quite plausible. To avoid it we employ exactly the same set of parameters along all the experiments.

Table 9
Results from profitability comparison

	Daily returns			Weekly returns		
	Linear	ANN	Difference	Linear	ANN	Difference
Emerging stock markets	7.98	7.42	***0.58***	16.38	11.83	***4.58***
Developed Stock Markets	8.16	13.69	***4.54***	11.25	15.70	***4.45***
Difference	***0.18***	***6.27***		***5.13***	***3.70***	

In this table we show the predictability of developed and emerging markets in economic terms (mean annualized returns) for daily and weekly horizons. The last row shows the difference, in terms of returns, between emerging and developed stock markets, so a positive number means more predictability in emerging markets, and vice versa. The third column, in each horizon, shows the difference in predictability between linear and neural network models, so a positive number indicates a better behavior of linear models against neural network.
Linear and Neural network models.

alternatives using out-of-sample data, choose the best model and then employ it in a real situation. For daily data, the sizes of the training, validation and testing windows are 250, 60 and 60, respectively (approximately one year for building the model, one quarter to evaluate it and another quarter of real trading). For weekly data they are 100, 25 and 25, respectively (two years for building the model, one semester to evaluate it and another semester of real trading).[21]

In Table 9 we show the results (in annualized terms) from the strategy explained above, to the extent that these average returns were higher than zero (or even, higher than the buy-and-hold strategy), one could affirm that some predictability exists in these stock markets. As we can see, all the results are negative, so we can conclude that from an economic perspective (and taking into account the transaction cost from each stock market), neither emerging stock markets nor developed ones are predictable.

Note, however, that there are some differences between daily and weekly data. We find that daily data is significantly more predictable than weekly data, in accordance with the results found in some other sections of this paper. As it is shown in Table 9 annualized losses from a linear model and daily frequency are, approximately, 50% and 25% lower for emerging and developed stock markets, respectively, than from a weekly frequency. In the case of ANN, the annualized losses from a daily frequency are 35% and 20% (for emerging and developed stock markets, respectively) lower than in weekly data. Therefore, we can conclude that even though returns seem to be unpredictable, they are even more difficult to forecast with data of a smaller frequency.

We can also observe how the ANNs do not have a clearly better behavior than linear models. The nonlinear models seem to outperform the linear ones only for emerging stock markets. Taking into account the conclusions from the last sections about the worse forecast capability of ANNs, we can conclude that arguing in favor of a superior behavior of these models when forecasting stock returns must be taken with caution.

In Table 10 we provide the results obtained along the horizon employed to build the model (under the column *mean training*), to validate it (*mean validation*) and of real trading (*mean test*), for weekly data.[22] We can observe how, for weekly data,[23] only two developed stock markets (Ireland and Sweden in which the net provides an annual profit equal to 0.9% and 4%, respectively) and two emerging ones (Czech Republic and Colombia) seem to be predictable. We must highlight the high annual profit (21.5%) obtained by the net in the case of Colombia but also the losses obtained in the case of some other emerging markets as Indonesia, Malaysia, and Greece, where the annual losses are higher than 30%. Again, these results let us to conclude that employing sophisticated models do not provide a clear advantage.

Although, in previous sections we have found that, in general, when lagged information from

[21] The experiments of this section are extremely computationally intensive, and forced us to restrict the number of configurations.

[22] Results for daily data are very similar. We do not present them to save space, but they are available upon request.

[23] For the daily case, the achieved return from all developed stock market is negative, and only it is positive for seven emerging stock markets (Indonesia, Korea, Philippines, Thai land, Chile, Peru, and Czech Republic).

Table 10
Loss/profit of artificial neural networks models, weekly returns

	Mean training (100 obs.)	Mean validation (25 obs.)	Mean test (25 obs.)
Emerging markets			
IND	0.012	0.068	0.220
INO	0.210	0.171	0.303
KOR	0.066	0.155	0.166
MAL	0.339	0.632	0.496
PHI	0.020	0.184	0.016
TAW	0.194	0.230	0.110
THA	0.046	0.290	0.274
ARG	0.202	0.010	0.036
BRA	0.368	0.631	0.215
CHE	0.095	0.160	0.004
COL	0.081	0.230	0.128
MEX	0.232	0.011	0.014
PER	0.018	0.068	0.022
VEN	0.407	0.201	0.126
CZE	0.031	0.125	0.020
GRE	0.446	0.064	0.363
HUN	0.596	0.108	0.208
SOU	0.039	0.150	0.207
TUR	0.420	0.004	0.044
Mean	**0.177**	**0.165**	**0.118**
Developed markets			
AUT	0.062	0.008	0.083
BEL	0.300	0.048	0.191
CAN	0.171	0.058	0.230
DEN	0.295	0.046	0.172
FIN	0.617	0.199	0.341
FRA	0.281	0.057	0.159
HON	0.125	0.031	0.277
IRE	0.247	0.138	0.009
ITA	0.445	0.099	0.228
JAP	0.034	0.017	0.243
LUX	0.035	0.022	0.020
NET	0.322	0.091	0.139
NOR	0.057	0.039	0.153
POR	0.438	0.101	0.094
SIP	0.079	0.164	0.280
SPA	0.567	0.071	0.114
SWE	0.299	0.043	0.040
SWI	0.333	0.051	0.331
UKG	0.269	0.097	0.075
USA	0.380	0.072	0.064
Mean	**0.258**	**0.034**	**0.157**

In this table we provide the mean annualized return achieved on each stock market following a trading strategy directed by an ANN. We provide the results obtained along the horizon employed to build the model (under the column *mean training*), to validate it (*mean validation*) and of real trading (*mean test*). We could not do the experiment for nine emerging stock markets (China, Pakistan, Sri Lanka, Egypt, Israel, Jordan, Morocco, Poland and Russia) because the transaction costs were not available.

USA is incorporated, both emerging and specially developed stock markets are more predictable, if equity transaction costs are considered predictability disappears. For example, Harvey et al. (2000) conclude that ANNs are a very useful tool to forecast emerging stock markets and they suggest that it would be possible to achieve an extra-profit over a buy-and-hold strategy. Specifically, in their study, they highlight the case of the Korean stock market and using weekly data from June 1997 to March 1999, they obtain 79% of earnings. According to our results (Table 10) when transaction cost are introduced we can conclude that there is not any extra-profit from investing in Korea, a result that can be attributable to the fact that Korea is one of the countries with the highest one-way transaction cost.[24] As a conclusion, we can affirm that the consideration of transaction costs is indispensable to evaluate the eventual predictability of both developed and emerging stock markets.

4. Conclusions

In this paper, we have analyzed the daily and weekly forecastability of stock returns of a large number of markets and several years. We employed different information sets as well as model specifications; we also considered linear and nonlinear forecasts to assess the validity of our results. Our results suggest that nonlinear models do not provide superior predictions than the linear ones and that emerging and developed stock markets are generally nonpredictable when total transaction cost are considered. Moreover, our study was carried out from two different perspectives: Firstly, from a statistical point of view, and secondly, from an economic perspective by taking into account trading costs.

From a statistical point of view we found some gains in predictability when one incorporates in the model the information produced by the USA market. We also found that the inclusion of information of other leading markets (Germany, Japan and the United Kingdom) do not allow us to obtain more accurate forecasts than univariate models. In addition, we found that developed markets seem to be slightly more predictable than the emerging ones. This may be explained by the fact that local information seems to be much more relevant for emerging markets than for developed markets, as already pointed out in a different context some authors (e.g. Rouwenhorst, 1999).

In economic terms, we showed that if total transaction costs are considered the results change completely: neither emerging stock markets nor developed ones show any degree of predictability. When compared with some other studies, our different results could be explained for two reasons: firstly, some papers have considered only explicit transaction costs, however, in many stock markets (especially emerging ones) the implicit costs are even higher, so not considering the whole trading cost could have lead to erroneous conclusions. Secondly, since we were interested in evaluating the economic significance of predictions, we employed a new loss function consistent with this objective. Since we optimized the nets using a measure of profit instead of the standard mean squared error, this change could modify the results obtained by the nonlinear models. As a conclusion, we can affirm that considering transaction costs is extremely relevant for predictability and, consequently, asset allocation decisions in both developed and emerging stock markets.

[24] According to Domowitz et al. (2001), although, the Korean explicit one way transaction cost (fees and commission) are not very high (63 basic points), its implicit cost (the difference between bid and ask prices, among others) are the highest ones (134.4 basic points) from the 42 stock markets considered.

References

Achour, D., Harvey, C.R., Hopkins, G., Lang, C., 1998. Stock selection in emerging markets: Portfolio strategies for Malay sia, Mexico and South Africa. Emerging Markets Quarterly Winter, 38 91.

Achour, D., Harvey, C.R., Hopkins, G., Lang, C., 1999. Stock Selection in Malaysia. Emerging Markets Quarterly (Spring), 54 91.

Ang, A., Bekaert, G., 2001. Stock return predictability: Is it there? NBER Working Papers, number 8207.

Barry, C.B., Goldreyer, E., Lockwood, L.J., Rodriguez, M., 2002. Robustness of size and value effects in emerging equity markets, 1985 2000. Emerging Markets Review 3 (1), 1 30.

Bekaert, G., Claude, B.E., Cambell, C.R., Viskanta, T.E., 1997. What matters for emerging equity market investments. Emerging Markets Quarterly (Summer), 17 46.

Bekaert, G., Harvey, C.R., Lundblad, C.T., 2003. Equity market liberalization in emerging markets. Journal of Financial Research 26 (3), 275 299.

Breen, W., Glosten, L.R., Jagannathan, R., 1990. Predictable variations in stock index returns. Journal of Finance 44, 1177 1189.

Brock, W.A., Dechert, W.D., Scheinkman, J.A., LeBaron, B., 1996. A test for independence based on the correlation dimension. Econometric Reviews 15 (3), 197 236.

Campbell, J.Y., 1987. Stock returns and the term structure. Journal of Financial Economics, 373 399.

Chang, E.J., Araújo Lima, E.J., Tabak, B.M., 2004. Testing for predictability in emerging equity markets. Emerging Markets Review 5 (3), 295 316.

Cheng, B., Titerington, D.M., 1994. Neural networks: A review from a statistical perspective. Statistical Science 9 (1), 2 54.

Clare, A.D., Thomas, S.H., Wickens, M.R., 1994. Is the gilt equity yield ratio useful for predicting UK stock returns? Economic Journal 104, 303 315.

Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253 265.

Domowitz, I., Glen, J., Madhavan, A., 2001. Liquidity, volatility and equity trading costs across countries and over time. International Finance 4 (2), 221 255.

Fama, E.F., French, K.R., 1998. Value versus growth: The international Evidence. Journal of Finance 53, 1975 1999.
Guo, H., 2006. On the out of sample predictability of stock market returns. Journal of Business 79 (2), 645 670.
Harvey, C.R., 1995. Predictable risk and returns in emerging markets. Review of Financial Studies 8, 773 816.
Harvey, C.R., Costa, Michael J., Travers, Kirsten E., 2000. Forecasting emerging market returns using neutral networks. Emerging Markets Quarterly 4 (2, Summer), 43 54.
Hornik, K., 1991. Approximation capabilities of multilayer feed forward networks. Neural Networks 4, 251 257.
Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359 366.
Hsieh, D., 1991. Chaos and nonlinear dynamics: Application to financial markets. Journal of Finance 46, 1839 1877.
Jasic, T., Wood, D., 2004. The profitability of daily stock market indices trades based on Neural Network predictions: Case study for the S&P500, the DAX, the TOPIX and the FTSE in the period 1965 1999. Applied Financial Economics 14 (4), 285 297.
Kohers, T.G., Pandey, V., 1998. The contribution of emerging markets in international diversification strategies. Applied Financial Economics 8 (5), 445 454.
Kuan, C.M., White, H., 1994. Artificial neural networks: An econometric perspective. Econometric Reviews 13, 1 91.
Levy, Haim, Sarnat, Marshall, 1970. International diversification of investment portfolios. American Economic Review 60 (4), 668 675.
Lewellen, J., 2004. Predicting returns with financial ratios. Journal of Financial Economics 74 (2), 209 235.
Milobedzki, P., 2004. Predictability of stock markets with disequilibrium trading. A commentary paper. European Journal of Finance 10 (5), 345 352.
Nelly, C.J., Weller, P., 2000. Predictability in international asset returns: A reexamination. Journal of Financial and Quanti tative Analysis 35 (4), 602 620.
Pesaran, M.H., Timmermann, A., 1994. Forecasting stock returns. An examination of stock market trading in the presence of transaction costs. Journal of Forecasting 13, 330 365.
Pesaran, M.H., Timmermann, A., 1995. The robustness and economic significance of predictability of stock returns. Journal of Finance 50, 1201 1228.
Pesaran, M.H., Timmermann, A., 2000. Recursive modelling approach to predicting UK stock returns. Economic Journal 110, 159 191.
Robbins, H., Monro, S., 1951. A stochastic approximation method. Annals of Mathematical Statistics 22, 400 407.
Rouwenhorst, K.G., 1999. Local return factors and turnover in emerging stock markets. Journal of Finance 54 (4), 1439 1464.
Solnik, B.H., 1995. Why not diversify internationally rather than domestically? Financial Analyst Journal 51 (1), 89 94.
Urrutia, J.L., 1995. Test of random walk and market efficiency for Latin American emerging equity markets. Journal of Financial Research 18 (Fall, 3), 299 310.
van der Hart, J., Slagter, E., van Dijk, D., 2003. Stock selection strategies in emerging markets. Journal of Empirical Finance 10 (1 2), 105 132.
Vose, M.D., 1999. The Simple Genetic Algorithm. The MIT Press.
White, H., 1989. Learning in artificial neural networks: A statistical perspective. Neural Computation 1, 425 464.
Zhang, G., Patuwo, B.E., Hu, M.Y., 1998. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting 14, 35 62.