



Weak-Form Market Efficiency of the Brazilian Exchange Rate: Evidence from an Artificial Neural Network Model

Andreza Aparecida Palma & Alexandre Sartoris

To cite this article: Andreza Aparecida Palma & Alexandre Sartoris (2016) Weak-Form Market Efficiency of the Brazilian Exchange Rate: Evidence from an Artificial Neural Network Model, Latin American Business Review, 17:2, 163-176, DOI: [10.1080/10978526.2016.1171720](https://doi.org/10.1080/10978526.2016.1171720)

To link to this article: <https://doi.org/10.1080/10978526.2016.1171720>



Published online: 30 Jun 2016.



Submit your article to this journal [↗](#)



Article views: 97



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Weak-Form Market Efficiency of the Brazilian Exchange Rate: Evidence from an Artificial Neural Network Model

Andreza Aparecida Palma^a and Alexandre Sartoris^b

^aDepartment of Economics, UFSCAR Federal University of São Carlos, Sorocaba, Brazil; ^bDepartment of Economics, UNESP São Paulo State University, Araraquara, Brazil

ABSTRACT

This article utilizes an artificial neural network model to examine the hypothesis of weak-form market efficiency for the monthly Brazilian exchange rate from 1999 to 2013. The method of partial derivatives suggested by Racine and White (2001) is used. The first step is to choose network architecture, second, weights estimation, and, at last, testing according to the suggested procedure. The results suggest that the Brazilian foreign exchange market is not efficient informally; thus, agents can obtain unusual profits through arbitrage.

RESUMEN

Mediante el uso del modelo de una red neural artificial, este estudio examina la hipótesis de una eficiencia débil en el mercado, para un tipo de cambio mensual brasileño entre 1999 a 2013. Para lograrlo, se utilizará el método de derivadas parciales sugerido por Racine y White (2001). El primer paso consiste en escoger la arquitectura de la red, seguido por valoraciones de los pesos y, finalmente, se realizará una prueba de conformidad con el procedimiento sugerido. Los resultados sugieren que el mercado de cambio brasileño no es eficiente (forma débil) y, consecuentemente, permite que se obtengan ganancias anormales a través del arbitraje.

RESUMO

O presente trabalho utiliza redes neurais artificiais para testar a hipótese de eficiência fraca de mercado para a taxa de câmbio mensal brasileira de 1999 a 2013. Para tanto, será usado o método das derivadas parciais, sugerido por Racine e White (2001). A primeira etapa deste método consiste na escolha da arquitetura da rede, seguida pelo treinamento e, por fim, do teste segundo o procedimento sugerido. Os resultados obtidos apontam que o mercado de câmbio brasileiro não é informalmente eficiente e, dessa forma, os agentes podem obter lucros anormais através de arbitragem.

ARTICLE HISTORY

Received 19 December 2015
Revised 29 January 2016
Accepted 29 January 2016

KEYWORDS

Artificial neural networks;
exchange rate; inference;
market efficiency

Introduction

Artificial neural networks (AANs), originally conceived by computer science and engineering as a tool of artificial intelligence, inspired in the

CONTACT Andreza Aparecida Palma ✉ drepalma@gmail.com 📠 Universidade Federal de São Carlos, Depto de Economia, Rodovia João Leme dos Santos, Km 110, Bairro do Itinga, Sorocaba, SP, CEP. 18052-780, Brazil. Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/wlab.

© 2016 Taylor & Francis

mechanism of the human brain, when used to evaluate the relationship between two or more variables in an empirical model, can be seen as nonparametric estimation methods, since they are not concerned with the data generator process and possess some advantages over traditional methods. Such advantages include the fact that it is not required to specify a functional form for the relationship among variables and its great capability of modeling nonlinear relationships. Because most economic relations are probably nonlinear (Granger, 1991 in Shachmurove, 2002, p. 7), the use of ANNs in economics has grown substantially in recent years. However, ANNs have certain disadvantages in comparison with traditional econometric models. One of them is the lack of a standard methodology for building these models, which are developed, therefore, on a trial and error basis (Shachmurove, 2002, pp. 27–28). Additionally, an ANN model is considered akin to a “black box” since the output generator process may not be explicitly known and thus cannot explain “how” and “why” such output was generated. Therefore, the estimated weights do not have a clear meaning, in contrast to, for instance, the coefficients in a regression model.

One of the disadvantages often mentioned regarding ANNs has been the lack of an existing method for doing formal statistical inference with those models (see, e.g., Zhang, Hu, & Patuwo, 1998, p. 56). However, Racine and White (2001) have shown that it is possible to use ANNs, more precisely, multilayer perceptron (MLP) networks to accomplish formal inference using bootstrap techniques.¹ Doing so, it is possible to carry out hypothesis tests of the individual and joint significance of the inputs in an ANN. Since a MLP network has the capability of extracting complex nonlinear relationships, inference based on such models becomes a very powerful tool of analysis. In this work, we will use an ANN model to test the weak-form efficiency hypothesis in the Brazilian foreign exchange market—we examine whether past values of the exchange rate have no influence on its current value.²

The empirical literature regarding market efficiency in Latin American countries is still scarce, despite the importance of these countries in international financial markets. Specifically, for Brazil, the empirical evidence is not conclusive. For example, Laurini and Portugal (2004) and Silva, Matsushita, Gleria, and Figueiredo (2007) found evidence of weak informational efficiency in the Brazilian exchange rate, while Tabak and Lima (2009) found evidence of medium- and long-term predictability. Further, according to Tabak and Lima (2009), the Brazilian foreign exchange market is worthy of analysis for several reasons. For example, in relation to developed economies, Brazil has higher exchange rate volatility. In addition, when compared to other emerging economies, Brazil has a well-developed derivatives market. Finally, the flexible exchange rate regime was adopted in Brazil only

in mid-1999 and may still be developing, suggesting some degree of inefficiency. Predictability analysis in an environment with these characteristics seems to be a major issue. Thus, this article contributes to the empirical literature in the field using a neural network model to formally test the weak-form efficiency hypothesis in the Brazilian foreign exchange market from 1999 to 2013.

The contribution of this study for the relevant empirical literature is twofold. First, we use a methodology based on ANNs (specifically, MLP), which takes into account the nonlinearity of the exchange rate. Second, we used a formal procedure to test the statistical significance of the lagged values of the Brazilian exchange rate in an ANN model. To our knowledge, this is the first study that uses a neuro-based approach to formally test the weak-form market efficiency of the Brazilian exchange rate. Note that the lack of a formal inference procedure is frequently cited as a drawback of ANNs. There is not yet a well-established approach to verify the significance of the ANNs inputs. Further, although there are many studies that use ANNs to compare the out-of-sample forecast performance,³ the focus of this article is to formally test the efficient-market hypothesis (weak-form), using the bootstrap procedure and partial derivatives suggested by Racine and White (2001).

The remainder of this study is organized as follows. In Section 2, we present the ANN model estimated for the Brazilian exchange rate. In Section 3, we conduct the statistical inference procedure to test the weak-form efficiency hypothesis. Section 4 is dedicated to final remarks.

An ANN for the nominal Brazilian exchange rate (R\$/US\$) – 1999–2013

Data description

We utilized monthly data from the nominal Brazilian exchange rate (purchase rate) in the period from February 1999 to December 2013, adding up 179 observations. The series were obtained from the website of the Central Bank of Brazil, <http://www.bcb.gov.br>. These chosen data are from the period after devaluation of the Real in 1999, when Brazil adopted a floating exchange rate regime.

Following Racine and White (2001, p. 666), in order to avoid nonstationarity issues, we use the logarithm of the first difference of the exchange rate—percent differences, or still, the returns of the exchange rate. Figure 1 shows the graph of the transformed series. It should be noticed that a MLP network requires a stationary series.⁴ The KPSS, ADF, and Phillips Perron tests were performed. All tests confirm that the series is stationary at 5% significance.⁵

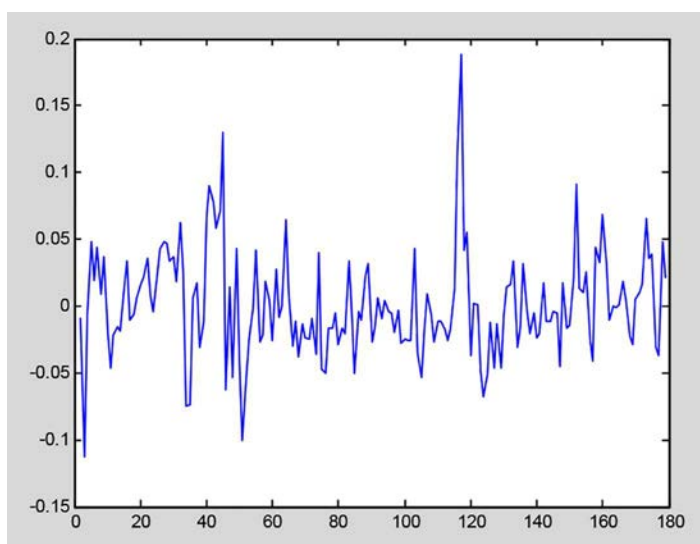


Figure 1. Brazilian exchange rate (percent difference) Feb. 1999–Dec. 2013. *Source:* Central Bank of Brazil.

$$\Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1}) \cong \Delta\%y_t$$

We also tested the linearity of log-returns of Brazilian exchange rate using the tests proposed by Lee, White, and Granger (1993) and Teräsvirta, Lin, and Granger (1993). The results in Table 1 indicate that we can reject the null hypothesis of linearity at 10%. Thus, a nonlinear model seems to really be the most appropriate.

Network architecture

Defining the network architecture is complicated, since there still is no well-established methodology for doing so. As Zhang and colleagues (1998, p. 42) pointed out, the definition of such parameters strongly depends on which problem the ANN is being used for. Because it is often poorly understood, many regard the process of choosing the architecture of an ANN as “black magic.” Some authors, such as Anders and Korn (1999), Medeiros and Teräsvirta (2001), Pizarro, Guerrero, and Galindo (2000), Ripley (1995), among others, suggest statistical procedures for this task. In this work, the choice of architecture followed, in essence, the procedure adopted by Racine and White (2001). We show such process next.

Table 1. Linearity test.

Test	Statistic	<i>P</i> -value
Terasvirta	5.9767	0.05037
White	6.2151	0.04471

Source: Calculated by the authors.

Number of inputs

The number of inputs, that is, the number of “independent variables,” is the number of lags of the exchange rate (or, more precisely, its percent variation). Since we are interested in testing the efficient-market hypothesis (in its weak version), we must check whether those lags are significant. In order to do so, we train networks using 1 to 10 inputs (i.e., 1 to 10 lags). The best one is chosen by the Schwarz Information Criterion (SIC).

Number of hidden layers

An ANN without a hidden layer is equivalent to a linear regression model. Since we are assuming nonlinearity for the exchange rate,⁶ it makes sense that we use at least one hidden layer. There are theoretical works showing that one hidden layer is enough to approximate any nonlinear function (see, e.g., Hornik, Stinchcombe, & White, 1989). Most works using ANNs for time series forecasting use only one hidden layer.⁷ According to Zhang and colleagues (1998, p. 44), one hidden layer seems to be suitable for most forecasting problems. Therefore, we are going to use networks with only one hidden layer.

Number of neurons in the hidden layer (Hidden nodes)

Usually, the number of neurons in the hidden layer is defined empirically. If we use too many neurons, the network could excessively memorize the training data, leading to an over fitting, and, as consequence, a low generalization capability. If we use a small number of neurons, the training could be extremely slow.

Following Racine and White (2001), the complexity of the network will be determined by using the SIC. In order to do so, we are going to use 1 to 10 hidden units and 10 random restarts of initial weights to each set of input units.⁸ The configuration that leads to the smallest value for the SIC, for each network, will be the chosen one, given the number of inputs.

Number of output neurons

That is the easiest part in defining the network configuration and the only one that is uncontroversial. Since we are interested in forecasting the exchange rate (its rate of return, actually) one-step ahead, we should use only one output neuron.

Activation function

The sigmoid (logistic) function⁹ is going to be used for the hidden layer neurons and the “pure” linear function for the output neuron. First, the best networks are chosen for each input configuration. Among those 10, the “best” network is chosen according to SIC. That is the one to be tested at the end of the work.

Network training and evaluation

We use the conjugate gradient algorithm for the training of the ANNs. One relevant issue when the training starts is whether the variables are going to be “normalized.”¹⁰ If so, it still remains to be decided which type of “normalization” is to be used. Even in our case, where the variables are normalized in some way (since all belong to the same time series), the normalization could still bring benefits (Azoff, 1994, p. 24). There is still no consensus about the need for normalization of the variables. Shanker, Hu, and Hung (1996), for instance, analyzing the benefits in network training by doing linear and statistical normalization of the variables.¹¹ Their conclusion was that there are, as a general result, benefits in performing some normalization; but those benefits decrease as both the sample and the network size increase.

Here we choose to train the networks using non-normalized variables and with statistical normalization, in order to make a comparison of the results. By analyzing the square correlation coefficient (R^2) between the network outputs and the targets (here called R^{2*}), the training using statistical normalization shows superior results.

It should be noticed that both the inputs and the targets were normalized. After the training, the outputs were “denormalized,” in order to analyze the results, as suggested by Zhang and colleagues (1998, p. 50). The chosen networks for each input configuration are shown in Table 2.¹²

Now, among those 10 networks, we should choose the “best.” As can be seen in Table 2, the ANN with one input was the one with smallest Schwarz criterion. Therefore, the chosen architecture is $1 \times 2 \times 1$. We also will perform the inference procedure considering a $10 \times 2 \times 1$ architecture, which showed better performance within the sample.¹³ In the next section, we perform statistical inference on these ANNs in order to test the validity of the efficient-market hypothesis in its weak version.

Table 2. “Best” ANN chosen for each configuration of the input layer.

Network architecture	Mean square error	Schwarz information criterion	ρ correlation coefficient	R^{2*}
$10 \times 2 \times 1$	0.000737702	−6.449476	0.712664	50.7889%
$9 \times 2 \times 1$	0.000753	−6.493137	0.706472	49.9103%
$8 \times 2 \times 1$	0.0007829	−6.518036	0.690747	47.7132%
$7 \times 1 \times 1$	0.001123	−6.49090	0.499273	24.9273%
$6 \times 2 \times 1$	0.0008741	−6.533476	0.642216	41.2442%
$5 \times 1 \times 1$	0.013630572	−4.057137	0.487371	23.7531%
$4 \times 1 \times 1$	0.013719996	−4.081353	0.478604	22.9062%
$3 \times 1 \times 1$	0.0140301	−4.089468	0.462045	21.3486%
$2 \times 2 \times 1$	0.012657	−4.105142	0.53534	28.6589%
$1 \times 2 \times 1$	0.0012929	−6.586738	0.52202	27.2505%

Source: Calculated by the authors.

Statistical inference in the ANN

In this section, hypothesis tests about the network inputs significance is performed using a bootstrap technique; but before that, a brief explanation of the efficient-market hypothesis is appropriate.

The efficient-market hypothesis

The idea of market efficiency is an important notion in the field of finance, and was originally articulated in the theoretical contribution of Bachelier (1900) who, in his doctoral thesis on “Speculation Theory,” mathematically developed the concept of the Brownian Motion.¹⁴ He began his work saying “past, present and even discounted future events are reflected in market price, but often show no apparent relation to price changes” (Bachelier in Dimson & Mussavian, 1998, p. 92) and concluded that commodities prices float randomly—the expected value for the speculator is zero. Bachelier was the first one to formulate reports and tests of the random walk model. Later studies showed that other financial series have the same feature. Working (1934), for instance, also identified the random behavior in prices of goods.¹⁵

In the fifties, Kendall (1953) began the systematization of the random walk theory. He did not manage to identify regular cycles of prices, concluding that they follow a random walk “in series of prices which are observed at fairly close intervals the random changes from one term to the next are so large as to swamp any systematic effect which may be present. The data behave almost like wandering series.” (Kendall, 1953 in Dimson & Mussavian, 1998, p. 92). Spurred by the development of computers, studies on asset prices grew rapidly. Roberts (1959) compared series generated by random numbers and current prices and concluded they are indistinguishable. Osborne (1959), analyzing US stock prices, showed they move in a way analogous to molecules. In the sixties, Samuelson (1965) and Mandelbrot (1966) mathematically formalized and “translated” into economic models the efficient market hypothesis.

The efficient-market hypothesis (EMH) has developed jointly with the Martingale model and, more specifically, with the random walk model. The Martingale model is basically a random walk, except for some constraints in the error term (Laurini & Portugal, 2004).

The efficiency concept is used to describe the fact that asset prices fully reflect all available information, thus conveying suitable signs for capital allocation. Thus, the best forecast for the current price is the past price, since it has in itself all the available information. Therefore, if the market is efficient, the changes in prices cannot be systematically forecast. The EMH implies, therefore, the nonexistence of any regularity or behavior pattern in the market that would allow abnormal profits.

In the EMH literature, there are three forms of efficiency (proposed initially by Roberts, 1967 and used by Fama, 1970), which are distinguished by the meaning of “relevant information,”—by the information set considered:

1. Weak efficiency: the information set available is in the history of the series, i.e., in its past values.
2. Semi-strong efficiency: the information set consists of all available public information. Thus, the market instantly absorbs all the information as soon as it becomes publicly available.
3. Strong efficiency: the information set is made of all available public and private information, i.e., also considering privileged information. Therefore, no agent could have abnormal earnings, even with access to confidential data.

Several empirical studies of the EMH for the Brazilian foreign exchange market have been undertaken and rather more for other markets, especially the stock market. Laurini and Portugal (2004) verified the validity of the EMH in its weak form, using the daily nominal exchange rate from January 7, 1994 to April 1, 2002. By using a Markov Switching model, they concluded that for some regimes in the series, the EMH is valid, although not in others. Garcia (2003) tested the efficiency of the term market for the exchange rate in Brazil also using daily data (January 18, 1999 to August 20, 2002). His results suggested that this market is not efficient. Tabak and Lima (2009) also found evidence of predictability in the Brazilian foreign exchange market.

In this work our interest is to test the weak form of EMH for the exchange market. To do so, we use the ANN chosen in the previous section ($1 \times 2 \times 1$ and $10 \times 2 \times 1$) and verify whether the past values (lags) of the exchange rate are significant when it comes to forecasting its present value. Usually, EMH testing is performed assuming a linear structure to the series. According to Bruni and Famá (1998, p. 76), the ability for forecasting depends mainly on the success of the nonlinear models. Since ANNs do have the capability of extracting nonlinear relationships without specification of a functional form, we will probably have a better result for the test made in an ANN than in a traditional model.

Valid statistical inference in ANNs: Tests based in partial derivatives

Consider the following MLP model with one hidden layer:

$$f(x, w) = w_{00} + \sum_{j=1}^h w_{0j} \psi(\tilde{x}' w_{1j})$$

where:

ψ = activation function of the network (logistic, in this case);

$$\tilde{x} = (1, x^T)^T$$

To test the hypothesis that some network inputs have no effect in the output is equivalent to testing the hypothesis that the partial derivatives from these inputs are zero, that is:

$$\frac{\partial f(x, w^*)}{\partial x_i} = 0 \quad i \in I_0 \quad (\text{I})$$

where:

w^* are the optimum weights found in the network training;

I_0 is the inputs set whose relevance we wish to test.

We could rewrite (I):

$$m^* = \sum_{i \in I_0} \int f_i(x, w^*)^2 d\mu(x) \quad (\text{II})$$

where:

$$f_i(x, w^*) = \frac{\partial f(x, w^*)}{\partial x_i}$$

$\mu(x)$ = probability distribution of X_t .

It follows that (I) will be true if, and only if $m^* = 0$. Notice, however, that w^* and $\mu(x)$ are unknown, but the weight \hat{w}_n found in the network training and the empirical distribution of X are consistently estimated by w^* and $\hat{\mu}_n$, respectively (Racine & White, 2001, p. 659). One feasible statistic is, therefore:

$$\hat{m}_n = \frac{1}{n} \sum_{t=1}^n \sum_{i \in I_0} f_i(X_t, \hat{w}_n)^2 = \int \sum_{i \in I_0} f_i(x, \hat{w}_n)^2 d\hat{\mu}_n(x).$$

For sufficiently large samples, Racine and White (2001, p. 660) showed that:

$$n\hat{m}_n \rightarrow N_2(0, C^*; M^*).$$

This distribution is not found in tables,¹⁶ but we can approximate it by using the bootstrap technique. The bootstrap statistics is given by:

$$\bar{\beta}_n^* \equiv \sum_{t=1}^n m(X_t, \hat{w}_n^*) - \sum_{t=1}^n m(X_t, \hat{w}_n) - \sum_{t=1}^n \nabla^T m(X_t, \hat{w}_n)(\hat{w}_n^* - \hat{w}_n)$$

where:

\hat{w}_n^* = optimum network weights trained with the resample of X_t and Y_t .

$$m(x, w) = \sum_{i \in I_0} f_i(x, w)^2.$$

And we have that (Racine & White, 2001, p. 662):

$$\bar{\beta}_n^* \xrightarrow{d} N_2(0, C^*; M^*).$$

We are interested in testing the hypothesis that past values of the exchange rate are not relevant to explaining future value. Thus, for each estimated model, we wish to test the joint hypothesis that all inputs are not relevant—the partial derivatives of $f(x, w^*)$ from the inputs are zero. This is done by following the steps described next.

1. A resample is computed with reposition (i.e., we take a new sample from the sample itself) from $\{X_t, Y_t\}$, and call it $\{X_t^*, Y_t^*\}$. We then estimate a new model with $\{X_t^*, Y_t^*\}$. The training algorithm will be initialized from the initial values \hat{w}_n found for the network trained before. With those resampled weights and the initial ones, we can therefore compute the bootstrap statistics:

$$\bar{\beta}_n^* \equiv \sum_{t=1}^n m(X_t, \hat{w}_n^*) - \sum_{t=1}^n m(X_t, \hat{w}_n) - \sum_{t=1}^n \nabla^T m(X_t, \hat{w}_n)(\hat{w}_n^* - \hat{w}_n).$$

2. The procedure described in (1) is repeated 1000 times, thereby obtaining 1000 values for the bootstrap statistics. We then compute the acceptance region of the null hypothesis (c_α).
3. The $n\hat{m}_n$ statistics will be computed. If it is bigger than c_α , then the null hypothesis of Market efficiency should be rejected.

Results

The results of the procedure described in the previous section are shown in Table 3.

As the results in Table 3 show, the null hypothesis that the network input variables - the lags of the nominal exchange rate - are not relevant, can be rejected, at a significance level of 10%. We could conclude, therefore, that the Brazilian exchange market is not weak-form efficient; thus, there is a possibility of abnormal earnings for the agents in this market.

Final remarks

In this work an application of the methodology suggested by Racine and White (2001) was performed to obtain formal statistical inference based on

Table 3. Empirical significance of network inputs $1 \times 2 \times 1$.

Statistics	Critical value (5%)	P-value
99.8603	104.3316	0.05300

Source: Calculated by the authors.

Table 4. Empirical significance of network inputs $10 \times 2 \times 1$.

Statistics	Critical value (5%)	P-value
65.848292	41.040108	0.036000

Source: Calculated by the authors.

ANN models. More precisely, we use that tool to verify the validity of the EMH for the Brazilian exchange rate market in the period from 1999 to 2013. The use of an ANN instead of a more traditional linear model is explained for the possibility of nonlinearity in the exchange rate series. The results of this work do not support the hypothesis of weak-form market efficiency for the Brazilian exchange rate for the period in question.

Although there are a reasonable number of studies in market efficiency for other markets in Brazil (e.g., stocks), there are only a few studies of market efficiency for the exchange rate. For example, Laurini and Portugal (2004) found evidence of weak informational efficiency in the Brazilian exchange rate, while Tabak and Lima (2009) found evidence of predictability in the medium and long term. According to Timmermann and Granger (2004, p. 15), attempts to forecast asset returns usually have little success, which supports the EMH. As pointed out by Granger and Teräsvirta (1993, p. 1), it is usually accepted that the economy is nonlinear and most economic variables have nonlinear relationships. Thus, the little success found so far for forecasting financial markets could be due to the use of essentially linear methodologies. Nonlinear models have seen substantial development since the eighties, but are still in early stages (Bruni & Famá, 1998, p. 76).

The use of nonlinear methodologies for testing EMH is therefore rather useful. The use of ANNs is particularly suitable since there is no need to specify a functional form among the variables. However, one of the disadvantages is the lack of standardized methodologies that ensure the specification of an “optimal” architecture for the network. Another disadvantage was the impossibility of statistical inference in an ANN model. However, recent developments, such as the work of Racine and White (2001), have overcome this deficiency. In sum, this article contributes to the empirical literature in the field using a neural network model to formally test the weak-form efficiency hypothesis in the Brazilian foreign exchange market from 1999 to 2013. The results suggest that the Brazilian foreign exchange market is not weak-form efficient informally; thus, agents can obtain unusual profits through arbitrage.

Notes

1. The bootstrap technique consists in using the sample itself in order to obtain a description of the sample properties of the estimators, instead of using theoretical results (Greene, 2003, p. 924).
2. The weak-form market efficiency hypothesis is generally verified by comparison of the forecasts generated by the model out of sample and those generated by the random walk model. The approach we use is to test the unpredictability of the market directly (as done, e.g., by Racine & White, 2001), in order to ascertain whether the values of past exchange rate returns are significant.

3. For example, Coelho, Santos, and Costa Jr (2008) compared the forecasting performance of ANNs for the Brazilian exchange rate at different frequencies with some econometric models. The conclusion is that the ANNs have a better forecasting performance.
4. Note, however, that other network architectures allow the use of nonstationary variables, such as, the recurrent networks.
5. The table containing the results can be obtained on request to the authors.
6. In fact, nonlinear models seem to be more suitable for exchange rates. Several works show that when a linearity test is carried out on exchange rate data, almost always evidence of nonlinearity is found (Franses & Van Griensven, 1997, p. 1).
7. There are, however, works showing the benefits of using two hidden layers. See, for instance, Barron (1994).
8. The random restart of the weights is recommended, since the training of an ANN is very sensitive to the initial weights. Several initialization methods have been suggested, such as, the initialization based on the results of a linear regression (Chan, Lam, & Wong, 2000), among others.
9. As Zhang and colleagues (1998, p. 47) and Swingler (1996, p. 62) noted that the use of the logistic function in the hidden neurons is the most popular choice for ANNs applications.
10. Normalization, here, in ANNs terminology, means transformation of the original data.
11. Linear normalization: $\frac{y_t - y_{\text{mínimo}}}{y_{\text{máximo}} - y_{\text{mínimo}}}$; statistical normalization: $\frac{y_t - \bar{y}}{dp(y_t)}$.
12. The notation $q \times p \times l$ -y means that the network has q inputs, p neurons in the hidden layer, 1 neuron in the output layer, and y means the y -th random restart of the weights.
13. It should be noted that the correlation analysis between the product generated by the ANN and the output variable does not penalize the degrees of freedom.
14. Brownian Motion was described for the first time by the botanist Robert Brown, in 1827. The concept was later developed mathematically by Albert Einstein in 1905 in the field of Physics and by Norbert Wiener, in the field of Mathematics. Both were unaware of Bachelier's work.
15. See also Cowles and Jones (1937) and Cowles (1944).
16. If $Z \sim N(\mu, V)$, then the quadratic form $Z'KZ \sim N_2(\mu, V; K)$. The χ^2 distribution is, therefore, a special case of the N_2 distribution (White, 1994, p. 170).

References

- Anders, U., & Korn, O. (1999). Model selection in neural networks. *Neural Networks*, 12, 309–323. doi:10.1016/s0893-6080(98)00117-8
- Azoff, E. M. (1994). *Neural network time series forecasting of financial markets*. Chichester: John Wiley and Sons.
- Bachelier, L. (1900). *Théorie de la Spéculation*. Paris: Gauthier-Villars.
- Bachelier, L. (1964). Theory of speculation. In P. Cootner (Ed.), *The character of stock market prices* (pp. 17–78). Cambridge, MA: MIT Press.
- Barron, A. R. (1994). A comment on “Neural networks: A review from a statistical perspective.” *Statistical Science*, 9(1), 33–35. doi:10.1214/ss/1177010640
- Bruni, A. L., & Famá, R. (1998). Eficiência, previsibilidade dos preços e anomalias em mercados de capitais: Teoria e evidências. *Caderno de Pesquisas em Administração*, 1(7), 71–85.
- Chan, M.-C., Lam, C.-C., & Wong, C.-C. (2000). *Financial time series forecasting by neural network using conjugate gradient learning algorithm and multiple linear regression weight initialization*. Computing in Economics and Finance, Hong Kong, n.61, 2000.

- Coelho, L. S., Santos, A. A. P., & Costa, N. C. A., Jr. (2008, December). Podemos prever a taxa de cambio brasileira? Evidência empírica utilizando inteligência computacional e modelos econométricos. *Gestão & Produção*, 15(3), 635–647. doi:[10.1590/s0104-530x2008000300016](https://doi.org/10.1590/s0104-530x2008000300016)
- Cowles, A. (1944). Stock market forecasting. *Econometrica*, 12, 206–214. doi:[10.2307/1905433](https://doi.org/10.2307/1905433)
- Cowles, A., & Jones, H. (1937). Some a posteriori probabilities in stock market action. *Econometrica*, 5, 280–294. doi:[10.2307/1905515](https://doi.org/10.2307/1905515)
- Dimson, E., & Mussavian, M. (1998). A brief history of market efficiency. *European Financial Management*, 4(1), 91–193. doi:[10.1111/1468-036x.00056](https://doi.org/10.1111/1468-036x.00056)
- Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25, 383–417. doi:[10.2307/2325486](https://doi.org/10.2307/2325486)
- Fama, E. (1970). The behavior of stock market prices. *The Journal of Business*, 38, 34–105. doi:[10.1086/294743](https://doi.org/10.1086/294743)
- Franses, P. H., & Van Griensven, K. (1997). Forecasting exchange rates using neural networks for technical trading rules. *Studies in Nonlinear Dynamics & Econometrics*, 2(4), 109–114. doi:[10.2202/1558-3708.1033](https://doi.org/10.2202/1558-3708.1033)
- Garcia, G. M. (2003). *Eficiência do mercado implícito de câmbio a termo no Brasil* (Dissertação de Mestrado). Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, Brasil.
- Granger, C. W. J. (1991). Developments in the nonlinear analysis of economic series. *Scandinavian Journal of Economics*, 93(2), 263–276.
- Granger, C. W. J., & Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. Oxford, UK: Oxford University Press.
- Hornik, K., Stinchcombe, S. M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2, 359–366. doi:[10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Kendall, M. (1953). The analysis of economic time series. *Journal of the Royal Statistical Society, Series A*, 96, 11–25.
- Laurini, M. P., & Portugal, M. S. (2004). Long memory in the R\$/US\$ exchange rate: A robust analysis. *Brazilian Review of Econometrics*, 24, 109–147.
- Lee, T. H., White, H., & Granger, C. (1993). Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests. *Journal of Econometrics*, 56(3), 269–290.
- Mandelbrot, B. (1966). Forecasts of future prices, unbiased markets, and Martingale models. *The Journal of Business*, 39, 242–255. doi:[10.1086/294850](https://doi.org/10.1086/294850)
- Medeiros, M. C., & Teräsvirta, T. (2001). Statistical methods for modelling neural networks. *Intelligent Systems*, 9:227–235.
- Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research*, 7, 145–173. doi:[10.1287/opre.7.2.145](https://doi.org/10.1287/opre.7.2.145)
- Pizarro, J., Guerrero, E., & Galindo, P. L. (2000). A statistical model selection strategy applied to neural networks. *Proceedings ESANN, 2000*, 55–60.
- Racine, J., & White, H. (2001). Statistical inference, the bootstrap, and neural network modeling with application to foreign exchange rates. *IEEE Transactions on Neural Networks*, 12(4), 1–19.
- Ripley, B. D. (1995). *Statistical ideas for selecting network architectures*. In B. Kappen & S. Gielen (Eds.), *Neural networks: Artificial intelligence and industrial applications* (pp. 183–190). London: Springer.
- Roberts, H. (1959). Stock market patterns and financial analysis: Methodological suggestions. *The Journal of Finance*, 14, 1–10. doi:[10.1111/j.1540-6261.1959.tb00481.x](https://doi.org/10.1111/j.1540-6261.1959.tb00481.x)
- Roberts, H. (1967). Statistical versus clinical prediction of the stock market. [Manuscript not published]. In *Center for research in security prices*. University of Chicago.

- Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6, 41–19.
- Shachmurove, Y. (2002). *Applying artificial neural networks to business, economics and finance*. [CARESS Working Paper no 8]. University of Pennsylvania.
- Shanker, M., Hu, M. Y., & Hung, M. S. (1996). Effect of data standardization on neural network training. *Omega*, 24(4), 385–397. doi:[10.1016/0305-0483\(96\)00010-2](https://doi.org/10.1016/0305-0483(96)00010-2)
- Silva, S., Matsushita, R., Gleria, I., & Figueiredo, A. (2007). Hurst exponents, power laws, and efficiency in the Brazilian foreign exchange market. *Economics Bulletin*, 7(1), 1–11.
- Swingler, K. (1996). *Applying neural networks: A practical guide*. San Francisco, CA: Morgan Kaufmann.
- Tabak, B. M., & Lima, E. J. A. (2009). Market efficiency of Brazilian exchange rate: Evidence from variance ratio statistics and technical trading rules. *European Journal of Operational Research*, 194(3), 814–820. doi:[10.1016/j.ejor.2008.01.005](https://doi.org/10.1016/j.ejor.2008.01.005)
- Terasvirta, T., Lin, C.-F., & Granger, C. W. J. (1993). Power of the neural network linearity test. *Journal of Time Series Analysis*, 14:205–20.
- Timmermann, A., & Granger, C. W. J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20, 15–27. doi:[10.1016/s0169-2070\(03\)00012-8](https://doi.org/10.1016/s0169-2070(03)00012-8)
- White, H. (1994). *Estimation, inference, and specification analysis*. New York: Cambridge University Press.
- Working, H. (1934). A random difference series for use in the analysis of time series. *Journal of the American Statistical Association*, 29, 11–24. doi:[10.1080/01621459.1934.10502683](https://doi.org/10.1080/01621459.1934.10502683)
- Zhang, G., Hu, M. Y., & Patuwo, B. E. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14, 35–62.