Predicting currency crisis contagion from East Asia to Russia and Brazil: an artificial neural network approach

Raphaël Franck[a,*], Aurélien Schmied[b]

[a] Bar Ilan University, Department of Economics, 52900 Ramat Gan, Israel.

[b] Université Paris-XII, Laboratoire de Systèmes Aléatoires, 61 Avenue du Général de Gaulle, 94010 Créteil, France.

Abstract

Studies dealing with currency crisis prediction are often vulnerable to data mining and perform poorly when tested on out-of-sample data. This paper suggests an artificial neural network approach to predicting speculative attacks. The properties of the multilayer perceptron are used to develop a method for predicting currency crises. It is then tested whether the 1998 and 1999 speculative attacks in Russia and Brazil were predictable, given the then recent turmoil in East Asian countries. Overall, it appears that the multilayer perceptron does a better job at predicting currency crises than a logit model.

Keywords: Artificial Neural Networks, Contagion, Crisis Prediction, Currency Crisis.

JEL Classification: F31, C45.

---

* Corresponding author. Tel: +972-3-5318935. Fax: +972-3-5353180. E-mail: franckr@mail.biu.ac.il. (R.Franck).

We thank André Fourçans, Miriam Krausz, Abraham Lioui, Gilles Pagès, Hillel Rappaport and participants at the Latin American Economic Association Conference (Madrid), the 9th "Approche Connexionniste en Sciences Economiques et en Gestion" Conference (Boulogne-sur-Mer) and the Bar Ilan Economics Workshop for helpful comments. The usual disclaimer applies.

1. Introduction

On July 2nd, 1997, the Thai baht lost 17% against the U.S. dollar after several months of speculative pressure. In a fortnight, the Philippine peso and the Malaysian ringgit were also on a float. On July 11th, the Indonesian monetary authorities widened the bands of the rupiah[1]. The Singaporean dollar, which was formally on a float, did not however come under pressure before the second week of August. The Singaporean authorities then decided not to defend the currency: it had already lost 8% against the U.S. dollar by mid-September. Both Taiwan and South Korea were spared from contagious speculative attacks during the two following months. There were not any significant speculative pressures on the Taiwanese currency until early October[2]. But speculative attacks quickly compelled the Taiwanese authorities to abandon the fixed exchange rate system: they decided to let the currency float on October 20th. In Korea, a policy of gradual adjustment had allowed the won to lose only 8.4% against the U.S. dollar between July and the end of October. The won nevertheless plummeted by 25% during the sole month of November. The Korean authorities formally allowed the currency to float on December 16th. Speculative pressures on emerging markets and economies in transition did not stop. Russia was to let the ruble float on August 17th, 1998. The crawling peg of the Brazilian currency was to be abandoned on January 13th, 1999[3].

---

[1] The Indonesian monetary authorities formally allowed the float of the rupiah on August 14th, 1997.

[2] At the same time, the Hong Kong currency suffered from speculative pressures. The float was however avoided because of the willingness of Hong Kong monetary authorities to raise short term interest rates drastically and of the existence of a currency board. Hong Kong is thus omitted from this research so as to focus on East Asian countries that had to let the exchange rate go.

[3] See Kaminsky and Schmukler (1999), Radelet and Sachs (1998a, 1998b) and Corsetti et al. (1999a, 1999b) for differing analyses of the East Asian crisis. See Desai (2000) and Chiodo and Owyang (2002) for accounts of the Russian crisis and Franco (2000) on the Brazilian crisis.

This paper aims at determining whether the currency crises in Russia and Brazil were predictable, given the then recent turmoil in East Asian countries. It differs from related studies, such as Baig and Goldfajn (2001), Hernández and Valdés (2001), and Kaminsky and Reinhart (2000), which investigate the channels of contagion[4] from East Asia to Russia and Brazil but do not study whether the Russian and Brazilian currency crises could have been predicted.

Our study builds upon the researches dealing with currency crisis prediction, which have been surveyed by Kaminsky et al. (1998) and Berg et al. (1999). These surveys show that traditional studies on currency crisis prediction often fail to detect signs of speculative attacks. This lack of performance may be related to flaws in these studies. First, they use samples that range over ten years or more, which is questionable. Indeed, it might be argued that the 1994 Mexican currency crisis had an effect on the 1999 Brazilian currency crisis, but it seems doubtful that the 1982 Mexican currency crisis somehow entailed the 1994 Mexican currency crisis. It must therefore be acknowledged that the macroeconomic causes of currency crises are likely to vary over time, and as such, relevant prediction models should have small and updated data samples.

Second, traditional studies on currency crisis prediction rely on monthly, quarterly and even yearly data to predict currency crises. Such a method cannot assess changes in market sentiment on a daily or a weekly basis and may therefore have limited prediction ability. In our opinion, estimations should rely on daily financial indicators so as to single out times of speculative pressures and periods of tranquility on the foreign exchange market.

[4] Studies on currency crisis contagion distinguish between political, financial and trade channels of transmission, as Dornbusch et al (2000) and De Bandt and Hartmann (2001) show in their surveys.

Third, traditional studies on currency crisis prediction assume that there are linear relationships between variables, which is far from being established[5]. Yet, it must be acknowledged that nothing justifies the existence of nonlinear relations between variables. However, since Berg and Patillo (1999a, 1999b, 1999c) show that studies on predicting currency crises that rely on linear cross-section models such as Bussière and Mulder (2000), Radelet and Sachs (1998a), Sachs et al. (1996) and Tornell (1999), are vulnerable to data mining and often perform poorly out of sample, taking into account the possibility of non-linearity may enhance the ability of models to predict currency crises.

Hence, given the limitations of previous studies dealing with currency crisis prediction, we develop in this paper an alternative method that is based on Artificial Neural Networks (ANNs). We discuss how the main properties of ANNs circumvent the problems related to the prediction of speculative attacks. We then test whether the currency crises in Russia and Brazil that occurred in 1998 and 1999 were predictable, given the then recent turmoil in East Asian countries. We also examine whether ANNs would have rightly indicated that, at the same time, neither Argentina nor Chile had to let their exchange rate go. Finally, we compare the results of our ANN estimators to those obtained with a standard logit model so as to assess the relevance of our ANN-based method.

The remainder of this paper is planned as follows. Section 2 presents the main properties of ANNs. Section 3 describes the model. Section 4 presents the data. Section 5 describes the empirical analysis. Section 6 concludes.

[5] In this respect, the recent study by Pozo and Amuedo-Dorantes (2003) is an exception to the literature as it suggests avoiding traditional parametric assumptions in currency crisis prediction by using extreme value theory.

2. The main properties of artificial neural networks

Building on studies by Anderson (1995), Cheng and Titterington (1994), Kuan and White (1994) and White (1989a, 1989b), we present the main properties of ANNs and discuss how they may be applied to currency crisis prediction. Readers familiar with ANNs may skip this section altogether.

ANNs are non-parametric multivariate statistical models. They stem from the so-called "connexionnist" approach in cognitive sciences that builds on a biological description of the human brain in order to design systems that reproduce some of its features, notably the ability to learn and to perform complex tasks, as well as to assess and solve problems. White (1989a) considers that ANNs are "applications of well-known statistical models […] to a novel class of nonlinear regression models".

An ANN may be viewed as a collection of transfer functions that relate input variables *X*, located on an input layer, to an output variable *Y*, located on an output layer, so that *Y*=*f*(*X*), where *f* is the function that stands for the architecture of the network. Figure 1 provides an example of such a network. The input units, called "neurons", send signals towards *Y* over connections that either attenuate or amplify the signals, depending on the transfer function ψ.

[Insert Figure 1]

Many transfer functions have been employed in ANN research. The most popular are the logistic function $\psi_1 : \Re \to \Re$ such that

$$\forall\, x \in \Re,\ \psi_1(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

and the hyperbolic tangent function $\psi_2 : \Re \to \Re$ such that

$$\forall\, x \in \Re,\ \psi_2(x) = \tanh(x) \qquad (2)$$

with $\tanh(x) = \dfrac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$

ANNs are more efficient when one or several layers of intermediate units are integrated in the network. Input units send signals to these intermediate units located on one or several layers which are said to be “hidden”. Networks with one or several hidden layers are referred to as multilayer ANNs. Figure 2 shows an example of a multilayer ANN with one hidden layer.

[Insert Figure 2]

In our opinion, several reasons justify using one of the most common multilayer ANNs, the multilayer perceptron, in order to overcome problems in currency crisis prediction[6]. Indeed, it has been shown by Hornik et al. (1989) and Hornik (1991) that the multilayer perceptron is a “universal approximator”: it is able to approximate any class of functions at any desired degree of accuracy. It may hence provide a linear (nonlinear) estimator if the relationships between variables are linear (nonlinear). Therefore, the multilayer perceptron should theoretically perform better than traditional models of currency crisis prediction. However, in econometrics, there is no “silver bullet” for the problem of convergence to local versus global optima. As such, we cannot know for sure that a multilayer perceptron avoids data mining. Hence, it must be acknowledged that an ANN approach to currency crisis prediction requires repeated estimation and comparison of results in order to determine whether these results converge or not to various local optima.

Another reason for repeated estimation of ANNs relates to the problem of “overfitting”. This means that we may have too few training data for too many neurons, i.e., the degree of freedom may be too important. In such a situation where the ANN is “large”, it is completely adapted to the data set and does a poor job of generalization when tested on out-

[6] Rosenblatt (1958, 1962) first developed the perceptron. There are already some applications of neural networks to macroeconomics that focus on forecasting macroeconomic time series. Studies by Maasoumi et al. (1996),

of-sample data. Moreover, if the ANN is too "small", i.e., it has too few neurons, it also provides irrelevant results on out-of-sample data. Hence, repeated estimations are needed in order to obtain a good estimator.

Thus, the possibility of "overfitting" in ANNs must be taken into account when dealing with currency crisis estimators, which must be estimated on small and updated data samples as we have argued in the introduction. In this paper, we follow Duin (1993, 1995, 1996) who considers that a "small" multilayer perceptron provides a relevant estimator if it is estimated over a sample that contains a minimum number of forty points[7].

In this study, we use two types of multilayer perceptron: one with one hidden layer, the other with two. Most research in ANNs considers that a multilayer perceptron with one hidden layer is usually sufficient. Still, it must be acknowledged that a multilayer perceptron with two hidden layers theoretically better captures nonlinear relationships between variables than a multilayer perceptron with one hidden layer. However, with three or four hidden layers, the estimator would not be really be consistent because the degree of freedom would be too important relatively to the number of data.

3. The model

This paper develops an ANN model of currency crisis prediction that builds upon the multilayer perceptron.

We first define the multilayer perceptron with one hidden layer. It has $p \times T$ input units $\left\{X_t^1, X_{t-1}^1, \ldots, X_{t-T}^1, X_t^2 \ldots, X_{t-T}^p\right\}$ with p the number of parameters and T the number of lags. It

---

Stock and Watson (1996), Swanson and White (1997), Moshiri and Cameron (2000) provide mixed but promising results, even when non-linearity is not explicit.

[7] It must be acknowledged that the bootstrap method, just like the multilayer perceptron, may be used on small data samples. However, the bootstrap lacks the ANNs' other properties that make them useful in currency crisis prediction. See Efron and Tishbirani (1993) on the properties of the bootstrap.

has n neurons on its single hidden layer, and one output unit Y which is a scalar. A threshold neuron which has a constant input that is equal to 1, is also defined. It is meant to help the ANN resist "abnormal" input data, and as such, be less vulnerable to data mining.

The multilayer perceptron with one hidden layer has the following form

$$Y_t = \psi_3\left(\sum_{k=1}^{n}\beta_k\psi_1\left(\sum_{l=1}^{p}\sum_{i=1}^{T}\eta_{lik}X_{t-i}^{l}+\eta_{0kj}\right)+\beta_{0j}\right) \tag{3}$$

where $(\eta_{lik})_{l\in\{1...T\},i\in\{1...p\},k\in\{1...n\}}$ and $(\eta_{0k})_{k\in\{1...n\}}$ are the coefficients of the model on the input layer and $(\beta_k)_{k\in\{1...n\}}$ and (β_0) are the coefficients on the hidden layer with $\{(\beta_k)_{k\in\{1...n\}},(\eta_{lik})_{l\in\{1...T\},i\in\{1...p\},k\in\{1...n\}},(\beta_0),(\eta_{0k})_{k\in\{1...n\}}\}\in\Re^{n^3\times T\times p}$. The transfer functions of the ANN are ψ_1 and ψ_2 where ψ_1 is the logistic function defined in equation (1) and ψ_2 is a linear transfer function $\psi_3:\Re\to\Re$ used to obtain the scalar Y such that

$$\forall\, \mathrm{x}\in\Re,\ \psi_3(x)=x \tag{4}$$

In the same manner, we define the multilayer perceptron with two hidden layers. Its input and output layers are identical to those of the multilayer perceptron with one hidden layer. It however has two hidden layers: the first one comprises n neurons and the second, m neurons. It is hence described by the following equation

$$Y_t = \psi_3\left(\sum_{j=1}^{m}\alpha_j\psi_1\left(\sum_{k=1}^{n}\beta_{kj}\psi_2\left(\sum_{l=1}^{p}\sum_{i=1}^{T}\eta_{likj}X_{t-i}^{l}+\eta_{0kj}\right)+\beta_{0j}\right)+\alpha_0\right) \tag{5}$$

where $(\eta_{lik})_{l\in\{1...T\},i\in\{1...p\},k\in\{1...n\}}$ and $(\eta_{0k})_{k\in\{1...n\}}$ are the coefficients of the model on the input layer, $(\beta_k)_{k\in\{1...n\}}$ and (β_0) are the coefficients on the first hidden layer, $(\alpha_j)_{j\in\{1...m\}}$ and α_0 are the coefficients on the second hidden layer with

$\{(\alpha_j)_{j\in\{1...m\}},(\beta_{kj})_{k\in\{1...n\},j\in\{1...m\}},(\eta_{likj})_{l\in\{1...T\},i\in\{1...p\},k\in\{1...n\},j\in\{1...m\}},\alpha_0,(\beta_{0j})_{j\in\{1...m\}},(\eta_{0kj})_{k\in\{1...n\},j\in\{1...m\}}\}\in\Re^{m^5\times n^3\times T\times p}$.

The transfer functions ψ_1, ψ_2 and ψ_3 are those that have respectively been defined in equations (1), (2) and (4).

For both estimators, the variable Y takes the following values

- if Y=0, there are no speculative pressures on the foreign exchange market;
- if Y=1, there are speculative pressures on the foreign exchange market.

Thus, both models may be viewed as "ANN logit models". In order to test the relevance of our multilayer ANN estimators, we also estimate a logit model. We do this by acknowledging, in line with White (1989a), that the ANN with no hidden layer that we described in Figure 1 is identical to a logit model if its transfer function is the logistic function of equation (1).

4. The data

4.1 The data sets

We use daily financial data that include the exchange rate of each country comprised in this study vis-à-vis the U.S. dollar, stock price indexes, interbank and deposit rates. The complete list of data is given in Appendix A.

Note that most of these domestic data are originally expressed in their national currency. But ANNs require data to be expressed in the same currency in order to provide a consistent analysis. This paper uses the U.S dollar as its benchmark currency. No bias is introduced when all data are converted to U.S dollar because no major exchange rate variation occurred before the currency crisis that we study in this paper.

4.2 Country classification

Before Russia and Brazil gave up on their crawling peg, Thailand, The Philippines, Malaysia, Indonesia, Singapore, Taiwan and Korea also had to let their exchange rate go.

In order to choose the content of the estimation sample, it seems natural to follow a chronological order. The first four countries which gave up on a fixed exchange rate system or widened their currency's floating bands, i.e. Indonesia, Malaysia, the Philippines and

Thailand, are included in the estimation sample. The three East Asian countries that subsequently suffered from speculative attacks, i.e. Singapore, Taiwan and Korea, will be used as a benchmark of the estimators' relevance. Throughout the rest of the paper, we will refer to these three countries as the "validation sample".

4.3 Sub-period classification

We now distinguish the tranquility and speculative pressure periods. We have justified in the introduction the use of a smaller learning sample than in traditional studies on currency crisis prediction. As such, we cannot use crisis-dating schemes which distinguish speculative periods over long-time periods by identifying sharp changes in the exchange rate à la Frankel and Rose (1996), by computing a weighted average of exchange rates and reserves à la Kaminsky et al. (1998), or a weighted average of exchange rate, reserves and interest à la Eichengreen et al. (1996a, 1996b).

This paper follows Baig and Goldfajn (1999, 2001) in identifying a priori the speculative pressure periods: these are the three- and five-month periods before the fall in the exchange rate. This choice is motivated on econometric and economic grounds. On the one hand, a three-month data sample at least comprises forty points. Following Duin (1993, 1995, 1996), this is the smallest sample that may be used if a consistent ANN model is to be estimated. On the other hand, economic reasons justify the use of a five-month sample. From all accounts on the East Asian crisis, the speculative pressures in Thailand started in February 1997, i.e. five months before the fall of the baht. If a bigger sample is used, the analysis may not be grounded anymore.

The tranquility periods, which are used in this study as a benchmark, must be identified. It is decided that the tranquility periods are within the one-year period prior to the beginning of speculative pressures.

The tests on the Russian and Brazilian data both begin on January 1st, 1998 and respectively end on August 14th, 1998 and on January 11th, 1999, even though the devaluations occurred on August 17th, 1998 and on January 13th, 1999[8]. If this study took into account the last days before the crises, the estimators would unmistakably detect the devaluation. This would jeopardize the test and it is therefore better not to include the last days before the fall of the pegs in the data.

Furthermore, in order to assess the relevance of the estimators, we test them on countries that did not have to adopt a flexible exchange rate system immediately in the wake of the East Asian crisis. We choose to test the estimators on Argentina and Chile. The choice of these countries is not fortuitous. While Chile has so far withstood speculative attacks, Argentina abandoned its currency board in early January 2002, although it seemed in a better situation than Brazil in January 1999.

We choose to estimate our models on Argentina and Chile from January 1st, 1998 to January 30th, 1999. We do not carry on the estimations after January 1999 since the time lag between the estimation period for the estimators (July 1997 to November 1997) and the test period would be too important: this would cast doubt on the pertinence of the analyzes.

Table 1 sums up the country and sub-period classifications.

[Insert Table 1 Here]

5. Estimation procedure

The estimation procedure is carried out with the Neural Network Toolbox of MATLAB. Models are first estimated. The out-of-sample accuracy of these models is tested on East Asian countries, then on Russia, Brazil, Argentina and Chile.

[8] The 14 and 17 August 1998 were a Friday and a Monday. The 11 and 13 January 1999 were a Monday and a Wednesday.

5.1 Estimation and model selection

In this sub-section, we aim at finding different specifications of the model which fit the data in the estimation sample. We also estimate models where the "validation sample" is part of the estimation data so as to obtain additional benchmarks on the relevance of the estimators.

For each estimator, we search for the optimal number of neurons on each layer. We use two model selection criterions: the mean square error (MSE) and the gradient. Estimators must minimize the former and be greater than the latter. We arbitrarily set the MSE to 0.01 and the gradient to 1.10^{-6}. All ANNs are able to minimize the MSE before reaching the gradient because we can search for the optimal architecture of the ANNs, i.e., the optimal number of neurons, which fits our selection criterions. This is however not possible for the logit model, whose architecture is by assumption constrained by not having a hidden layer. Therefore, the MSE cannot be equal to 0.01 for the logit model. As such, the fitness of the logit on the estimation sample is always less satisfactory than that of the ANNs.

Furthermore, in order to estimate the models, we use the backpropagation algorithm, developed by Rumelhart et al. (1986), which is of a very frequent use in ANN research[9]. This algorithm may be presented as follows.

In an ANN, let ψ_i be the activation function of unit i. Let $U_i = \sum_l \theta_{li} X_l$ be the sum of units prior to unit i and $X_i = \psi_i(U_i)$. Let Y_{iz} and $\hat{Y}_{iz}$ be the i[th] components of the observed output and the computed output of the network that correspond to an entry *z* that is presented to the unit at each time period. The adaptation formulas of a connection that links unit *k* to unit *i* are

$$\Delta\theta_{ki} = \eta\delta_i X_k \tag{6}$$

where η is a strictly positive real number and

$$\delta_i = \left(Y_{iz} - \hat{Y}_{iz}\right)\psi_i'\left(U_i\right) \tag{7}$$

if unit i belongs to the output layer or

$$\delta_i = \left(\sum_{j=1}^{J} \delta_j \theta_{ij}\right)\psi_i'\left(U_i\right) \tag{8}$$

if unit i belongs to the other layers.

Table 2 gives the results of the training process for the various specifications[10]. Their relevance must now be tested with out-of sample data.

[Insert Table 2 here]

5.2 Out-of-sample accuracy: Singapore, Taiwan and South Korea

We now assess the accuracy and fitness of the estimators on out-of-sample data from Singapore, Taiwan and South Korea. These three countries, which make up our "validation sample", were hit by speculative attacks after July 1997.

Figure 3 shows the tranquility and speculative periods for the "validation sample". If the estimators obtained in the previous sub-section are relevant, we should obtain a graphic similar to Figure 3.

[Insert Figure 3 here]

---

[9] In this study, we could not estimate any models using algorithms that required an accurate computation of the gradient. Thus, we employed the backpropagation algorithm, because it avoids an accurate computation of the gradient. See Luenberger (1984) on this.

[10] From Table 2, it may seem that the estimators have too many parameters. However, when estimators with fewer neurons are trained, the MSE is far from 0.01. This means that models with fewer neurons are not helpful when it comes to our data sets.

Table 3 sums up the results of this test and assesses the out-of-sample performance of the various estimators by using two performance indicators, the Mean Absolute Error (MAE) and the Mean Square Error (MSE), which are defined as follows

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - T_t\right| \tag{9}$$

and

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - T_t\right)^2 \tag{10}$$

where Y_t is the value predicted by the estimator at time t, T_t is the target value, n the number of observations. In line with our assumptions concerning the estimation of the models, the target value T_t, $t=\{1,..n\}$, equals 1 for speculative pressure periods and 0 for tranquility periods for each of the three countries in our "validation sample".

[Insert Table 3 here]

Not surprisingly, models whose estimation data also include the “validation sample” perform well. They also perform better than models whose estimation data exclude the "validation sample". It must be noted that the performance of these latter models is not poor per se with respect to the MAE and MSE. However, the estimators fail to provide relevant crisis signal, as can be seen in Figure 4 which shows the results which are obtained with the one and two-layer ANNs, as well as with the logit, using the five-month estimation sample without the “validation sample”.

[Insert Figure 4 here]

Both the ANNs and the logit fail to provide good results for this test. There are several explanations for this failure. It may be that neither the ANNs nor the logit model are relevant currency crisis predictors. It may also be that our data sample lack significant variables. Finally, the lack of performance of the ANNs and the logit may stem from economic causes. Indeed, the specification that is used in this paper assumes that the Singaporean, Taiwanese and Korean crises are similar to those that occurred in Thailand, Indonesia, Malaysia and Philippines. But Singapore and Taiwan had relatively sound fundamentals before their currencies came under speculative pressure. What actually triggered market participants to launch attacks is their noticing that the Singaporean and Taiwanese currencies had become overvalued following the crises in the other East Asian countries, which were their main commercial partners. The situation was different in Korea which had been going through a severe economic crisis since the second semester of 1996. The Brazilian economy also suffered from macroeconomic and financial difficulties before July 1997. Moreover, the outbreak of the East Asian crisis surely increased the speculative pressures on the ruble and the real. It may then have been possible to predict the currency crises in Brazil.

5.3 Out-of sample accuracy: Russia, Brazil, Chile and Argentina

We now test the estimators on Russia, Brazil, Chile and Argentina data. Tables 4 and 5 sum up the results.

[Insert Table 4 here]

[Insert Table 5 here]

We first discuss the results obtained on data from Russia and Brazil, i.e., the two countries that had to let their exchange rate go. Concerning Russia, no multilayer perceptron with one hidden layer is able to identify speculative pressures. The two-hidden layer perceptron whose "validation sample" is not included in the estimation sample provide crisis signals throughout the test period, whether estimated on the three- or five-month sample. But so is the logit model estimated on the five-month sample that includes the "validation sample". Overall, the two-hidden layer perceptron and the logit do an equal job at predicting the Russian crisis.

As for Brazil, it appears that the bulk of estimators give crisis signals of the then upcoming speculative attack. Note however that the two-hidden layer ANN estimated over a three month sample without the "validation sample" provide unclear patterns for some time periods: they display brief alternating moments of tranquility and speculative pressure, which are impossible to interpret. As an example, we graph in Figure 5 the results of the one- and two-hidden layer perceptrons and of the logit, estimated on the five-month sample without the "validation sample". It appears that both ANNs give crisis signals throughout the period, while the logit only provides signals starting from the second half of 1998 until the crisis.

[Insert Figure 5]

Needless to say, our results, and as such the pertinence of our explanations, rest on the quality of our estimators. To assess their relevance, we provide an out-of-sample test on Argentina and Chile, i.e., two countries that did not let their exchange rate go between July 1997 and January 1999.

As for Argentinean data, it appears that the logit either does not give any crisis signals at all, or gives them throughout the test period. The multilayer perceptron with two hidden layers also provides poor results, by mostly providing false signals of speculative attacks. However, the multilayer perceptron with one hidden layer, which is estimated on the five-month sample with the "validation sample" included in the estimation sample, provides rather relevant signals. It indicates that Argentina did not suffer from speculative pressures, except for a couple of brief periods in the wake of the Brazilian currency crisis. In Figure 6, we graph this estimator, along with the multilayer perceptron with two hidden layers and the logit model estimated on the same data sample, so as provide an example of these models' relative performance.

[Insert Figure 6]

Concerning Chile, the multilayer perceptron with one hidden layer does not give evidence of any speculative pressure throughout the test period. However, the logit model provides some periods of speculative pressure, and so does the multilayer perceptron with two hidden layers. Given that the macroeconomic situation was relatively stable in Chile despite the financial turmoil in East Asia and Brazil, it seems that these crisis signals indicate the failure of the logit and of the multilayer perceptron with two hidden layers rather than actual crisis signals.

All in all, the quality of the results obtained from these out-of-sample tests on Argentina and Chile is mixed. The logit performs rather correctly on Argentina, but not Chile. The multilayer perceptron with two hidden layers fails to assess the lack of speculative pressures in both countries. Finally, the multilayer perceptron with one hidden layer, which provided satisfying results on the Brazilian data, performed rather well on both Argentina and Chile.

Furthermore, the fact that the multilayer perceptron with one hidden layer outperforms the logit for Brazil, Chile and Argentina may be seen as an indication that there were nonlinearities in the data sample that we used in this study to predict currency crises. Such an observation therefore provides an additional justification for using nonlinear models, rather than linear ones, in future studies on currency crisis prediction.

As an additional tool of comparison of the various estimators, we compute in Table 6 their MAE and MSE. In line with our assumptions concerning the estimation of the models, the target value T_t, $t=\{1,..n\}$ of the MAE and MSE in equations (9) and (10) equals 1 for Brazil and Russia and 0 for Argentina and Chile, since the two former countries let their currency float and the two latter did not.

[Insert Table 6 here]

In Table 6, it appears that the out-of-sample performances of the one- or two-hidden layer perceptrons as measured by the MAE and the MSE are superior to those of the logit for all countries except Russia. This is noteworthy as Russia is the only country where the perceptron does not provide better crisis signals than the logit. Such an observation suggests that perceptrons with low MAEs and MSEs are likely to be good crisis predictors.

6. Conclusion

In this study, we analyze the extent to which the spread of the East Asian crisis to Brazil was predictable using ANNs. For comparison purposes, we also examine the predictive ability of a logit model.

This study suggests that the fundamental determinants of currency crises in the first four East Asian countries that were affected by speculative attacks and in the last three that subsequently had to let the exchange rate go were different. It also implies that the speculative attacks in Thailand, Malaysia, the Philippines and Indonesia do not bear a strong resemblance to the Russian and Brazilian crises. This study also indicates that Argentina might have briefly suffered from the repercussions of the Brazilian crisis, but that Chile remained unaffected.

This study also casts a new light on currency crisis prediction. Berg and Patillo (1999a) already investigated whether it was possible to predict currency crises and answered: “Yes, but not very well”. In this paper, a type of ANNs, the multilayer perceptron with one hidden layer, is shown to provide relevant crisis signals for Brazil and to outperform a logit model on Argentina and Chile by not providing false indications of upcoming speculative attacks. Furthermore, ANNs provide a promising path of research because they are able to overcome problems usually associated with currency crisis prediction.

Appendix A

In this study, we use daily financial data from the Datastream software that are converted from each national currency into U.S. dollar. This is done by using the spot exchange rate at the opening of the market for each currency to the U.S. dollar as given by the Datastream software.

Stock market indexes: Dow Jones World Index, Dow Jones Bank Index, Dow Jones Basic Industries Index, Dow Jones Financial Institutions Index, Dow Jones Telecom Index and Dow Jones Utilities Index.

Interbank rate: Overnight rate.

Deposit rates: 90-day deposit rate, 180-day deposit rate.

References

Anderson, J.A., 1995, An introduction to neural networks (MIT Press, Cambridge, MA).

Baig, T., Goldfajn, I., 1999, Financial market contagion in the Asian crisis, International Monetary Fund Staff Papers 46 (2), 167-195.

Baig, T., Goldfajn, I., 2001, The Russian default and the contagion to Brazil. In S. Claessens and K.J. Forbes, eds., International financial contagion (Kluwer Academic, Dordrecht), 267-299.

Berg, A., Borensztein, E., Milesi-Feretti, G.M., Patillo, C., 1999, Anticipating balance-of-payments crises: the role of early warning systems, IMF Occasional Paper, (International Monetary Fund, Washington DC).

Berg, A., Patillo, C. 1999a, Are currency crises predictable?, a test, IMF Staff Papers 46 (2), 107-38.

Berg, A., Patillo, C. 1999b, Predicting currency crises: the indicators approach and an alternative, Journal of International Money and Finance 18 (4), 561-586.

Berg, A., Patillo, C. 1999c, What caused the Asian crises: an early warning system approach, Economic Notes 28 (3), 285-334.

Bussière, M., Mulder C., 2000, Political instability and economic vulnerability, International Journal of Finance and Economics 5 (4), 309-330.

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

Chiodo, A.J., Owyang, M.T., 2002, A case study of the currency crisis: the Russian default of 1998, Federal Reserve Bank of St. Louis 84 (6), 7-17.

Corsetti, G., Pesenti, P., Roubini, N., 1999a, Paper tigers? a model of the Asian crisis, European Economic Review, 43 (7), 1211-1236.

Corsetti, G., Pesenti, P., Roubini, N., 1999b, What caused the Asian currency and financial crisis?, Japan and the World Economy 11 (3), 305-373.

De Bandt, O., Hartmann, P. 2000, Systemic risk: a survey, CEPR Discussion Paper n°2634.

Desai, P., 2000, Why did the ruble collapse in August 1998 ?, AEA Papers and Proceedings 90 (2), 49-52.

Dornbusch, R., Park, Y.C., Claessens, S., 2000, Contagion: how it spreads and how it can be stopped, The World Bank Research Observer 15 (2), 177-197.

Duin, R.P.W., 1993, Superlearning capabilities of neural networks?, SCIA '93, Proceedings 8th of the Scandinavian conference on image analysis, Tromso, 547-554.

Duin, R.P.W., 1995, Small sample-sized generalizations, SCIA '95, Proceedings of the 9th Scandinavian conference on image analysis, Tromso, 957-964.

Duin, R.P.W., 1996, A note of comparing classifiers, Pattern Recognition Letters 17 (5), 529-536.

Eichengreen, B., Rose, A.K., Wyplosz, C., 1996a, Contagious currency crises, CEPR Discussion Paper n°1453.

Eichengreen, B., Rose, A.K., Wyplosz, C., 1996b, Exchange market mayhem: the antecedents and aftermath of speculative attacks, Economic Policy 21, 251-312.

Efron, B., Tishbirani, R.J., 1993, An introduction to the bootstrap (Chapman & Hall, Boca Raton, FLA).

Franco, G.H.B., 2000, The real plan and the exchange rate, Essays in International Finance 217, Princeton University.

Frankel, J., Rose, A.K., 1996, Currency crashes in emerging markets: an empirical treatment, Journal of International Economics 41, 351-366.

Hernández, L.F., Valdés, R.O., 2001, What drives contagion: trade, neighborhood or financial links?, International Review of Financial Analysis 10 (3), 203-218.

Hornik, K., 1991, Approximation capabilities of multilayer feedforward networks, Neural Networks 4, 251-257.

Hornik, K., Stinchombe, M., White, H., 1989, Multilayer feedforward networks are universal approximators, Neural Networks 2, 359-366.

Kaminsky, G.L., Lizondo, S., Reinhart, C.M., 1998, Leading indicators of currency crises, International Monetary Fund Staff Papers 45, 1-48.

Kaminsky, G.L., Reinhart, C.M., 2000, On crises, contagion, and confusion, Journal of International Economics 51 (1), 169-181

Kaminsky, G.L., Schmukler, S.L., 1999, What triggers market jitters? a chronicle of the Asian crisis, Journal of International Money and Finance 18, 537-560.

Kuan, C.M., White, H., 1994, Artificial neural networks: an econometric perspective, Econometric Reviews 13 (1), 1-91.

Luenberger, D.G., 1984, Linear and nonlinear programming (Addison-Wesley, Reading, MA).

Maasoumi, E., Khotanzad, A. Abaye, A., 1996, Artificial neural networks for some macroeconomic series: a first report, Econometric Reviews 13 (1), 105-22.

Moshiri, S., Cameron, N., 2000, Neural network versus econometric models in forecasting inflation, Journal of Forecasting 19, 201-217.

Pozo, S., Amuedo-Dorantes, C., 2003. Statistical distributions and the identification of currency crises, Journal of International Money and Finance 22, 591-609.

Radelet, S., Sachs, J., 1998a, The East Asian financial crisis: diagnosis, remedies, prospects, Brookings Papers on Economic Activity 1, 1-90.

Radelet, S., Sachs, J., 1998b, The onset of the East Asian crisis, NBER Working Paper 6680.

Rosenblatt, F. 1958, The perceptron: a probabilistic model for information storage and organization in the brain, Psychological Review 65, 386-408.

Rosenblatt, F. 1962, Principles of neurodynamics: perceptrons and the theory of brain mechanisms (Spartan, Washington DC).

Rumelhart, D.E, Hinton, G.E., Williams, R.J., 1986, Learning internal representations by error propagation, in: D.E. Rumelhart, J.L. McClelland and the PDP Research Group, eds., Parallel distributed processing, Vol. 1, (MIT Press, Cambridge, MA) 318-362.

Sachs, J., Tornell, A., Velasco, A., 1996, Financial crises in emerging markets: the lessons from 1995, Brookings Papers on Economic Activity 1, 147-215.

Stock, J.H., Watson, M.W., 1996, Evidence on structural instability in macroeconomic time series relations, Journal of Business and Economic Statistics 14 (1), 11-30.

Swanson, N., White, H., 1997, A model selection approach to real-time macroeconomic forecasting under linear models and artificial neural networks, The Review of Economic and Statistics 79, 540-550.

Tornell, A., 1999, Common fundamentals in the Tequila and Asian crises, NBER Working Paper 7139.

White, H., 1989a, Some asymptotic results for learning in single hidden-layer feedforward network models, Journal of the American Statistical Association 84, 1003-1013.

White, H., 1989b, Learning in artificial neural networks: a statistical perspective 87, 1252.

White, H., 1992, Some asymptotic results for learning in single hidden-layer feedforward network models (correction), Journal of the American Statistical Association, 87, 1252.

Table 1. Country and sub-period classification

	3-month data sample		5-month data sample	
	Beginning and end dates of the tranquility period	Beginning and end dates of the speculative pressure period	Beginning and end dates of the tranquility period	Beginning and end dates of the speculative pressure period
Learning sample				
Thailand	02-Apr-1996 - 02-Jul-1996	02-Apr-1997 - 02-Jul-1997	02-Feb-1996 - 02-Apr-1996	02-Feb-1997 - 02-Jul-1997
Indonesia	11-Apr-1996 - 11-Jul-1996	11-Apr-1997 - 11-Jul-1997	11-Feb-1996 - 11-Jul-1996	11-Feb-1997 - 11-Jul-1997
The Philippines	11-Apr-1996 - 11-Jul-1996	11-Apr-1997 - 11-Jul-1997	11-Feb-1996 - 11-Jul-1996	11-Feb-1997 - 11-Jul-1997
Malaysia	12-Apr-1996 - 15-Jul-1996	14-Apr-1997 - 14-Jul-1997	12-Feb-1996 - 15-Jul-1996	14-Feb-1997 - 14-Jul-1997
Validation sample – East Asian countries				
Singapore	10-May-1996 - 12-Aug-1996	12-May-1997 - 12-Aug-1997	12-Mar-1996 - 12-Aug-1996	12-Mar-1997 - 12-Aug-1997
Taiwan	19-Jul-1996 - 21-Oct-1996	18-Jul-1997 - 20-Oct-1997	20-May-1996 - 21-Oct-1996	20-May-1997 - 20-Oct-1997
South Korea	21-Aug-1996 - 21-Nov-1996	21-Aug-1997 - 21-Nov-1997	21-Jun-1996 - 21-Nov-1996	20-Jun-1997 - 21-Nov-1997
Out-of sample countries				
Russia	-	01-Jan-1998 - 14-Aug-1998	-	01-Jan-1998 - 14-Aug-1998
Brazil	-	01-Jan-1998 - 11-Jan-1999	-	01-Jan-1998 - 11-Jan-1999
Validation sample – Latin American countries				
Argentina		01-Jan-1998 - 30-Jan-1999		01-Jan-1998 - 30-Jan-1999
Chile		01-Jan-1998 - 30-Jan-1999		01-Jan-1998 - 30-Jan-1999

Table 2. Training specification and results

Estimator	Size of the sample	The validation sample is included in the estimation sample	Number of neurons of the first and second hidden layers	Mean Square Error	Gradient
1-layer ANN	3 months	No	7,0	0.0099307	0.048396
1-layer ANN	3 months	Yes	10,0	0.0099872	1.042684
1-layer ANN	5 months	No	7,0	0.00999835	0.0175472
1-layer ANN	5 months	Yes	12,0	0.00999765	0.106895
2-layer ANN	3 months	No	5,4	0.00998046	0.151932
2-layer ANN	3 months	Yes	5,4	0.00999853	0.0240339
2-layer ANN	5 months	No	5,4	0.00998311	0.37532
2-layer ANN	5 months	Yes	7,5	0.00999575	0.262232
Logit	3 months	No	0,0	0.138912	1.10^{-6}
Logit	3 months	Yes	0,0	0.240904	1.10^{-6}
Logit	5 months	No	0,0	0.20311	1.10^{-6}
Logit	5 months	Yes	0,0	0.233552	1.10^{-6}

Table 3. Training specification, simulation results, MAE and MSE of the estimators on Singaporean, South Korean and Taiwanese data.

Estimator	Size of the sample	The validation sample is included in the estimation sample	Do the simulation results fit the validation sample data ?	MAE Singapore	MSE Singapore	MAE South Korea	MSE South Korea	MAE Taiwan	MSE Taiwan
1-layer ANN	3 months	No	No	0.9771	0.9907	0.6856	0.6335	1.0312	1.0742
1-layer ANN	3 months	Yes	Yes	0.0190	7.6496e-004	0.0984	0.0150	0.0229	6.9235e-004
1-layer ANN	5 months	No	No	0.8403	0.9633	0.8073	0.7213	0.8350	0.9500
1-layer ANN	5 months	Yes	Yes	0.0209	5.0958e-004	0.0652	0.0074	0.0226	8.0591e-004
2-layer ANN	3 months	No	No	0.2664	0.1136	0.5516	0.5537	0.2943	0.1437
2-layer ANN	3 months	Yes	Yes	0.0109	2.2661e-004	0.1181	0.0645	0.0126	1.8189e-004
2-layer ANN	5 months	No	No	0.7643	0.7151	0.6613	0.6586	0.5227	0.3785
2-layer ANN	5 months	Yes	Yes	0.0351	0.0012	0.2023	0.0565	0.0350	0.0017
Logit	3 months	No	No	0.4898	0.2883	0.4622	0.4622	0.5111	0.3035
Logit	3 months	Yes	Yes	0.3995	0.1619	0.4802	0.2478	0.4114	0.1667
Logit	5 months	No	No	0.8403	0.9633	0.8073	0.7213	0.8350	0.9500
Logit	5 months	Yes	Yes	0.6061	0.3043	0.2962	0.1292	0.6164	0.3093

Table 4. Training simulation and results on Russian and Brazilian data

Estimator	Size of the sample	The validation sample is included in the estimation sample	Crisis signal in Russia	Crisis signal in Brazil
1-layer ANN	3 months	No	No	From 29-Jan-1998 to 21-Aug-1998
1-layer ANN	3 months	Yes	No	From 01-Jan-1998 to 13-Jan-1999
1-layer ANN	5 months	No	No	From 01-Jan-1998 to 13-Jan-1999
1-layer ANN	5 months	Yes	No	No
2-layer ANN	3 months	No	From 01-Jan-1998 to 14 Aug-1998	No[b]
2-layer ANN	3 months	Yes	No	From 11-Mar-1998 to 7-Sep-1998 and from 13-Nov-1998 to 08-Jan-1999[c]
2-layer ANN	5 months	No	From 01-Jan-1998 to 04-Jun-1998 and from 01-Jul-1998 to 07-Aug-1998[a]	From 01-Jan-1998 to 13-Jan-1999
2-layer ANN	5 months	Yes	No	From 02-Jan-1998 to 07-Sep-1998 and from 09-Nov-1998 to 13-Jan-1999
Logit	3 months	No	No	From 01-Jan-1998 to 13-Jan-1999
Logit	3 months	Yes	No	From 01-Jan-1998 to 13-Jan-1999
Logit	5 months	No	From 01-Jan-1998 to 14 Aug-1998	From 29-Jul-1998 to 08-Sep-1998, from 16-Sep-1998 to 08-Jan-1999 and on 13-Jan-1999
Logit	5 months	Yes	No	From 01-Jan-1998 to 13-Jan-1999

[a] Note: No clear pattern between 05-Jun-1998 and 6-Aug-1998

[b] Note: No clear pattern between 9-Sep-1998 and 12-Nov-1998

[c] Note: No clear pattern between 01-Jan-1998 and 10-Mar-1998

Table 5. Training simulation and results on Argentinean and Chilean data

Estimator	Size of the sample	The validation sample is included in the estimation sample	Crisis signal in Argentina	Crisis signal in Chile
1-layer ANN	3 months	No	No	No
1-layer ANN	3 months	Yes	From 26-Aug-1998 to 21-Sep-1998, from 29-Sep-1998 to 14-Oct-1998 and from 11-Jan-1999 to 30-Jan-1999	No
1-layer ANN	5 months	No	From 12-Dec-1998 to 06-Jan-1999	No
1-layer ANN	5 months	Yes	No	No
2-layer ANN	3 months	No	No	No clear pattern
2-layer ANN	3 months	Yes	No	From 01-Jan-1998 to 30-Jan-1999
2-layer ANN	5 months	No	From 01-Jan-1998 to 30-Jan-1999	From 01-Jan-1998 to 01-Apr-1998[b]
2-layer ANN	5 months	Yes	From 02-Mar-1998 to 05-Aug-1998[a]	No clear pattern
Logit	3 months	No	No	On 23-Aug-1998, 25-Aug-1998, 08-Sep-1998 and 10-Sep-1998
Logit	3 months	Yes	From 01-Jan-1998 to 30-Jan-1999	No
Logit	5 months	No	No	No
Logit	5 months	Yes	From 01-Jan-1998 to 30-Jan-1999	From 01-Jan-1998 to 30-Jan-1999

[a] No clear pattern between 01-Jan-1998 and 01-Mar-1998

[b] No clear pattern between 02-Apr-1998 and 30-Jan-1999

Table 6. MAE and MSE of the estimators on Argentinean, Brazilian, Chilean and Russian data.

Estimator	Size of the sample	The validation sample is included in the estimation sample	MAE Brazil	MSE Brazil	MAE Russia	MSE Russia
1-layer ANN	3 months	No	5.5411	74.3868	13.3297	177.9857
1-layer ANN	3 months	Yes	0.0402	0.0311	4.927	24.2845
1-layer ANN	5 months	No	0.0022	5.0331 10-6	12.1067	146.5730
1-layer ANN	5 months	Yes	2.7889	7.9876	1.5535	2.5333
2-layer ANN	3 months	No	0.9641	0.9369	0.2725	0.0745
2-layer ANN	3 months	Yes	0.3214	0.279	0.9331	0.8706
2-layer ANN	5 months	No	0.1026	0.0105	0.5192	0.4024
2-layer ANN	5 months	Yes	0.1394	0.1174	1.9006	1.3786
Logit	3 months	No	0.2866	0.0822	1	1
Logit	3 months	Yes	0.4115	0.1693	0.6012	0.3614
Logit	5 months	No	0.3886	0.2267	0.0591	0.0035
Logit	5 months	Yes	0.4685	0.2195	1.3625	1.8565

Estimator	Size of the sample	The validation sample is included in the estimation sample	MAE Argentina	MSE Argentina	MAE Chile	MSE Chile
1-layer ANN	3 months	No	1.1481	1.3486	5.337	29.3826
1-layer ANN	3 months	Yes	0.6954	0.5671	7.9952	63.9574
1-layer ANN	5 months	No	0.9773	1.0089	11.6332	135.4385
1-layer ANN	5 months	Yes	3.2811	10.7656	4.9576	24.6814
2-layer ANN	3 months	No	0.9728	0.9464	0.5896	0.5544
2-layer ANN	3 months	Yes	0.9334	0.8713	0.0399	0.0088
2-layer ANN	5 months	No	0.1025	0.0105	0.3488	0.1901
2-layer ANN	5 months	Yes	0.1394	1.0428	0.3798	0.2418
Logit	3 months	No	1	1	0.9873	0.9832
Logit	3 months	Yes	0.4115	0.1693	0.6012	0.3614
Logit	5 months	No	0.6188	0.3829	0.6188	0.3829
Logit	5 months	Yes	0.4685	0.2195	0.4685	0.2195

Fig. 1. The architecture of a two-layer ANN.

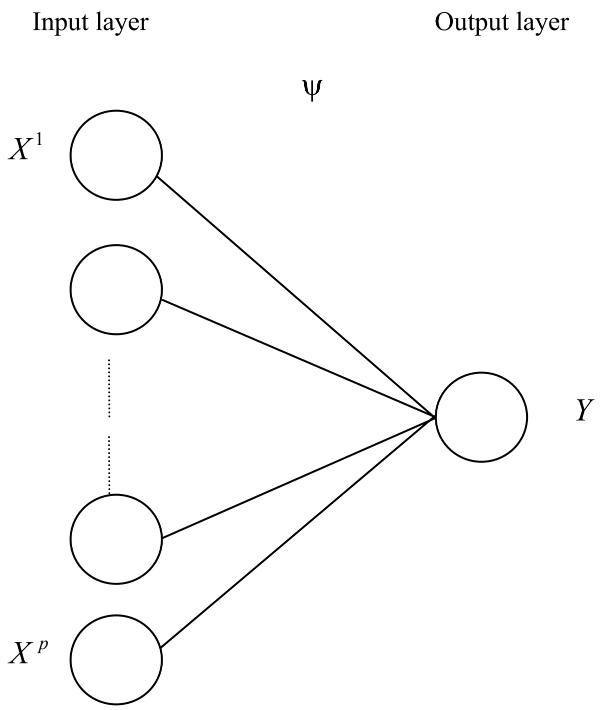

Fig. 2. The architecture of the multilayer perceptron with one hidden layer.

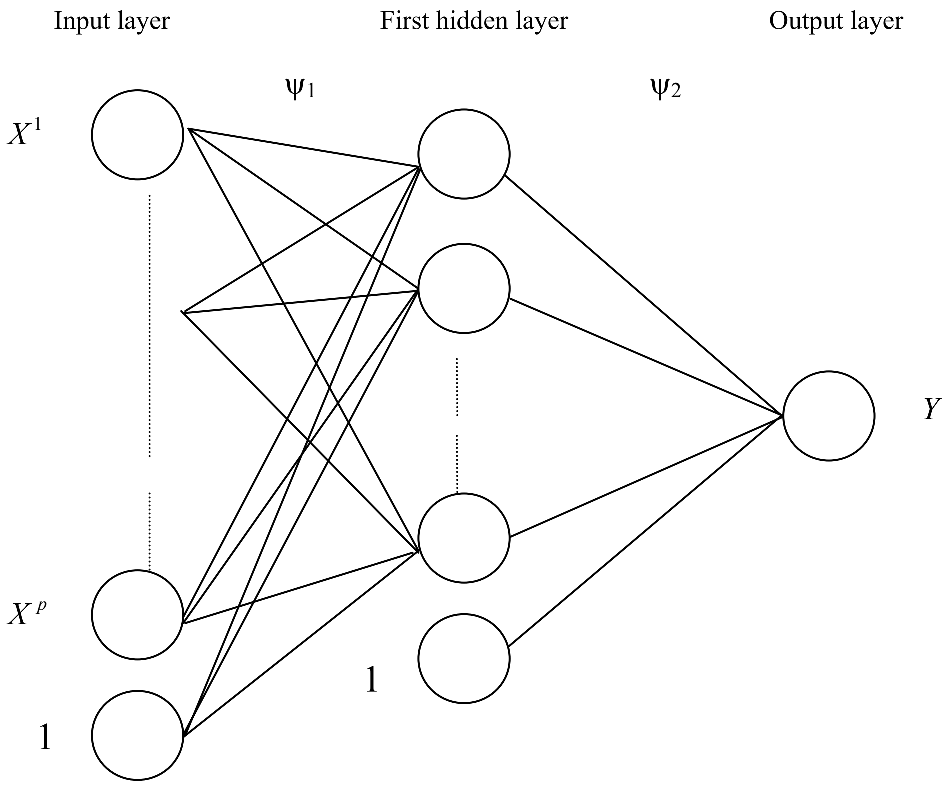

Fig. 3. The absolute validation data set (five-month sample).

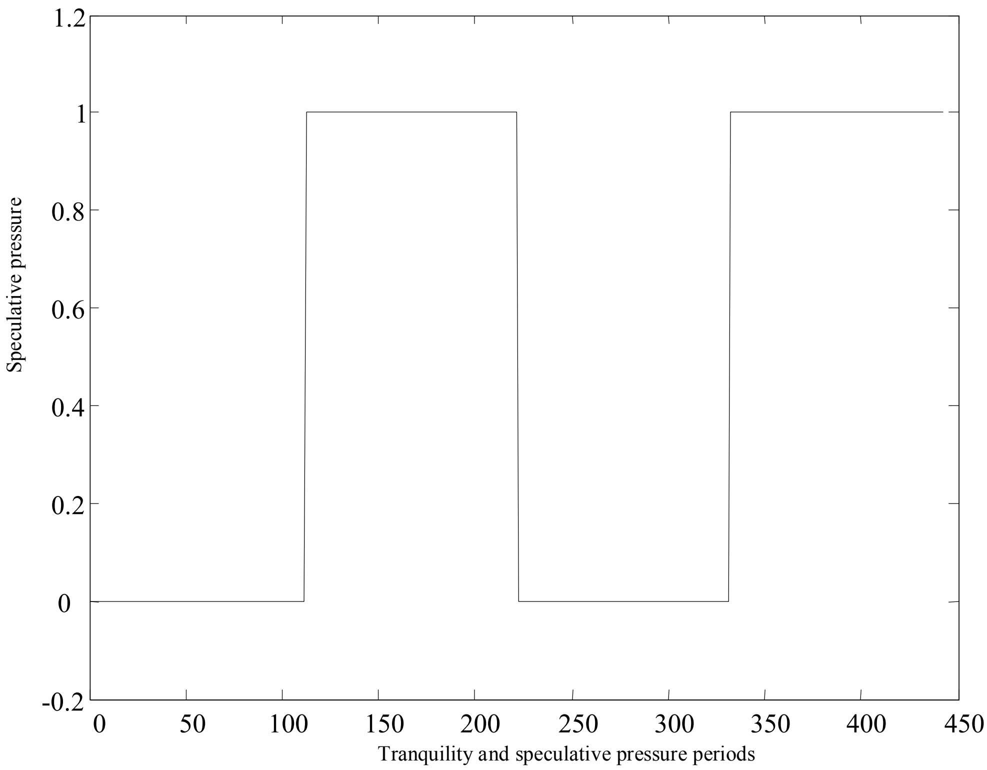

Fig. 4. Out-of sample accuracy of the five-month validation sample

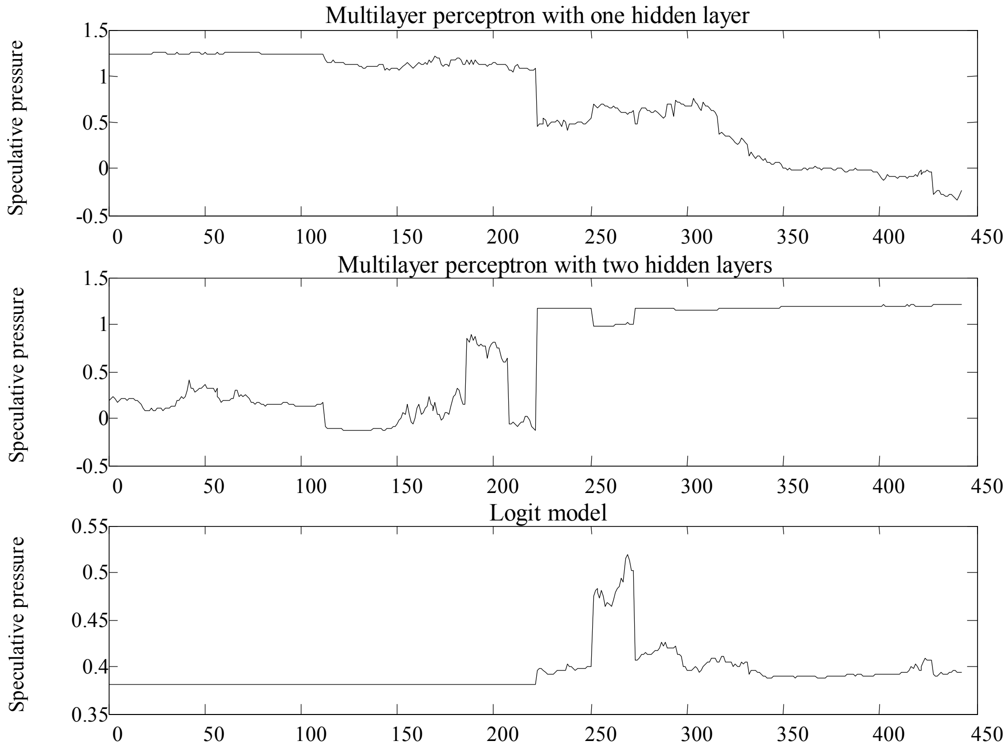

Fig. 5. Out-of sample accuracy on Brazilian data (five-month sample).

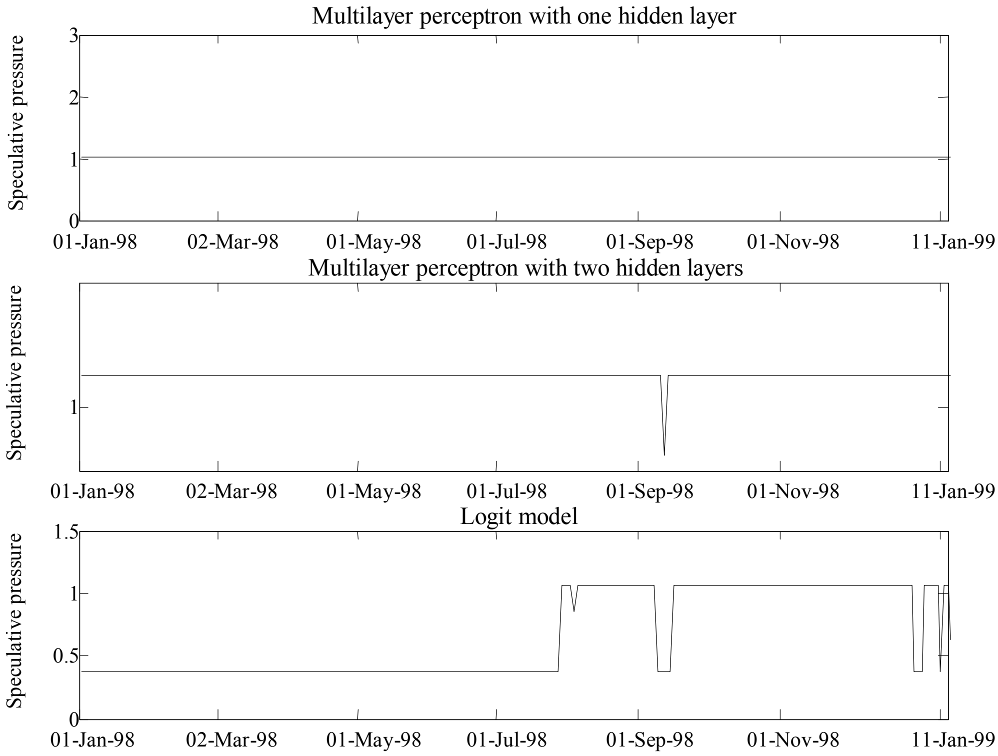

Fig. 6. Out-of sample accuracy on Argentinean data (five-month sample).