ISTITUTO DI STUDI E ANALISI ECONOMICA

Predicting Sovereign Debt Crises Using Artificial Neural Networks: A Comparative Approach

by

Marco Fioramanti
ISAE, piazza dell'Indipendenza, 4, 00185 Roma

Working paper n. 72
October 2006

The Series "*Documenti di Lavoro" of the Istituto di Studi e Analisi Economica* – Institute for Studies and Economic Analyses (ISAE) hosts the preliminary results of the research projects carried out within ISAE. The diffusion of the papers is subject to the favourable opinion of an anonymous referee, whom we would like to thank. The opinions expressed are merely the Authors' own and in no way involve the ISAE responsability.

The series is meant for experts and policy-makers with the aim of submitting proposals and raising suggestions and criticism.

La serie "Documenti di Lavoro" dell'Istituto di Studi e Analisi Economica ospita i risultati preliminari di ricerche predisposte all'interno dell'ISAE: La diffusione delle ricerche è autorizzata previo il parere favorevole di un anonimo esperto della materia che qui si ringrazia. Le opinioni espresse nei "Documenti di Lavoro" riflettono esclusivamente il pensiero degli autori e non impegnano la responsabilità dell'Ente.

La serie è destinata agli esperti e agli operatori di politica economica, al fine di formulare proposte e suscitare suggerimenti o critiche.

Stampato presso la sede dell'Istituto

ISAE - Piazza dell'Indipendenza, 4 – 00185 Roma.
Tel. +39-06444821; www.isae.it

ABSTRACT

Recent episodes of financial crises have revived the interest in developing models that are able to timely signal their occurrence. The literature has developed both parametric and non parametric models to predict these crises, the so called Early Warning Systems. Using data related to sovereign debt crises occurred in developing countries from 1980 to 2004, this paper shows that a further progress can be done applying a less developed non-parametric method, i.e. Artificial Neural Networks (ANN). Thanks to the high flexibility of neural networks and to the Universal Approximation Theorem an ANN based early warning system can, under certain conditions, outperform more consolidated methods.

Keywords: Early Warning System; Financial Crisis; Sovereign Debt Crises; Artificial Neural Network.

JEL Classification: F34; F37; C45; C14.

NON-TECHNICAL SUMMARY

Financial crises occurred in emerging countries in the last decade of 20th century have revived theoretical and empirical interest in the topic in order to understand their causes and consequences as well as to develop statistic and econometric models that can timely signal their occurrence. Economic theory has developed three generation of models explaining financial crises: the "first" and "second generation" models focus on currency crises and public imbalances, while "third generation" models include a wider variety of crises and are better suitable at explaining episodes occurred in the late '90s which were caused, principally, by private imbalances.

In the last decade, many empirical studies have concentrated their attention in developing models able to timely signal the occurrence of a financial crisis, the so-called early warning system (EWS). Using statistical and econometric techniques these models are applied to predict the likelihood of financial crises using a wide number of indicators related to internal and external factors, as well as social and political condition.

The aim of this paper is to further develop the Early Warning System literature related to financial crisis. In particular, a less explored non-parametric method, i.e. Artificial Neural Network (ANN), is carried to test its ability in predicting crisis imminence.

This paper, it confronts the performances of Artificial Neural Networks (ANNs) with those of a traditional parametric method: the Random Effect Probit estimator (REP). This paper shows, with an empirical application to sovereign debt crises in developing countries, that a well developed ANN can outperform both parametric and non-parametric traditional methods in timely signal crises episodes.

PREVISIONE DELLE CRISI DI DEBITO DI EMITTENTI SOVRANI ATTRAVERSO L'UTILIZZO DELLE RETI NEURALI ARTIFICIALI: UN APPROCCIO COMPARATO

SINTESI

Le recenti crisi finanziarie verificatesi nell'ultimo decennio hanno riacceso l'interesse verso lo sviluppo di modelli capaci di segnalare in anticipo il verificarsi di tali crisi. Nel corso degli anni la letteratura empirica relativa agli Early Warning System ha sviluppato modelli econometrici sia parametrici sia non parametrici. Usando i dati relativi alle crisi di debito sovrano verificatesi nei paesi emergenti nel periodo 1980-2004, questo articolo mostra che un ulteriore progresso può essere fatto nel perfezionare tali modelli utilizzando un metodo non parametrico fino ad ora poco sviluppato nella pratica: le Reti Neurali Artificiali (ANN). Grazie all'elevata flessibilità delle reti neurali e al Teorema di Approssimazione Universale, in questo articolo si mostra come un Early Warning System basato sulle reti neurali attenga, sotto determinate condizioni, performance migliori dei metodi maggiormente diffusi.

Parole chiave: Early Warning System; Financial Crisis; Sovereign Debt Crises; Artificial Neural Network.

Classificazione JEL: F34; F37; C45; C14.

CONTENTS

1 INTRODUCTION[1]

The aim of this paper is to further develop the Early Warning System literature related to financial crisis. In particular, a less explored non-parametric method, i.e. Artificial Neural Network (ANN), is carried to test its ability in predicting crisis imminence. This paper shows, with an empirical application to sovereign debt default in developing countries, that a well developed ANN can outperform both parametric and non-parametric traditional methods in timely signal crises episodes.

Financial crises occurred in emerging countries in the last decade of 20th century have revived theoretical and empirical interest in the topic in order to understand their causes and consequences as well as to develop statistic and econometric models that can timely signal their occurrence [2].

According to Krugman (1999, 2001) and Kaminsky (2003) economic theory has developed three generation of models explaining financial crises: the “first” and “second generation” models focus on currency crises, while “third generation” models include a wider variety of crises and are better suitable at explaining episodes occurred in the late ’90s. In the "first generation" models poor economic policies contrast with the goal of fixed exchange rate and produce a continuous loss in foreign exchange reserve. Once reserves fall below a critical level, authorities are forced to abandon the exchange rate peg. The building blocks of "second generation" models are the existence of multiple equilibria for exchange rate and self-fulfilling speculative attack. Even in the presence of sound economic policy the costs of maintaining a fixed exchange rate can be considered too high by government once the currency is subject to speculative attack. If investors doubt about authorities commitment to maintain the peg and start to sell the currency, government is induced to abandon the peg which would have been otherwise sustainable. In this sense speculative attacks are self-fulfilling. "Third generation" models were developed after the Asian crises and shifted the focus from public to private imbalances because public finances in those countries were quite sound, while those of corporate and banking sector showed excesses. The literature not only analyzes currency crises, but also bank and "twin crises" (currency and baking crises), balance of payment crises, and sovereign debt crises. The theoretical underpins of third generation are various: the moral hazard problem due to an implicit government

1 I wish to thank an anonymous referee and MIUR for financial support.

2 For a review of this topics see the book of Roubini and Setser (2005). Sturzenegger and Zettelmeyer (2006) concentrate their effort on sovereign default.

guarantee which, together with poor regulation, induces to over-landing and over-investment; the balance sheet effect due to the mismatch between assets and liabilities; self-fulfilling liquidity run when (government, bank or corporate) debt has short term maturity; sudden-stop to capital inflow due to external shocks. In all these cases the currency crisis is "more a symptom than a fundamental aspect of these crises"[3], and government, bank, corporate and currency crises are often related each other.

In the last decade, many empirical studies have concentrated their attention in developing models able to timely signal the occurrence of a financial crisis, the so-called Early Warning System (EWS). Using statistical and econometric techniques these models are applied to predict the likelihood of financial crises using a wide number of indicators related to internal and external factors, as well as social and political condition. According to the type of approach, models can be classified between parametric and non-parametric[4]. Frankel and Rose (1996) and Kaminsky et al. (1998) are the seminal papers in the two classes of approach applied to currency crises prediction. Using a probit, Frankel and Rose estimate the probability of currency crises for more than 100 developing countries from 1971 to 1992, finding that crises occur when GDP growth and the ratio of FDI to external debt are low, while growth of domestic credit and foreign interest rates are high. Kaminsky et al. proposed the non-parametric Signal Approach (or KLR) which involves monitoring the evolution of a number of economic indicators that show a behaviour which is different in tranquil period and prior to a crisis. When an indicator exceeds a particular threshold, this is interpreted as a signal that a crisis could occur in the following 24 months. Using a dataset of 23 countries from 1970 to 1995 and a wide set of indicators, they show that international reserves, real exchange rate, domestic credit, credit to public sector, and domestic inflation are very useful in signalling a crisis.

In parallel with the previous one, empirical literature on debt crises, which is rather small compared with that related to currency crises, can be classified between parametric and non-parametric. Detragiache and Spilimbergo (2001) focus on external debt and, using a probit analysis, try to test the hypothesis of liquidity run due to maturity mismatch. They use macroeconomic, debt and liquidity variables of developing countries to construct a model able to predict a crisis. The main finding is that liquidity variables show high significance, confirming the theory of self-fulfilling liquidity run. More recently Ciarlone and

---

[3] Krugman (2001) pg. 8.

[4] For a review of different approaches see Abiad (2003), Berg et al. (2004), and Ciarlone and Trebeschi (2004).

Trebeschi (2005) use a multinomial logit specification to predict entry into a crisis[5], and show that it outperform a simple logit. Their main result is that the level of external debt, international reserve, debt service and the degree of openness are very significant in predicting the probability of a crisis. Manasse et al. (2003) use both parametric (logit) and non-parametric (Classification and Regression Tree Analysis – CART) approach to develop an early warning system for debt crises. They show that a combination of the two approach improves the performance of the logit in predicting entry into a crisis.

This paper tries to further develop the application of non-parametric methods in predicting sovereign debt crises in developing countries. In particular, it confronts the performances of Artificial Neural Networks (ANNs) with those of a traditional parametric method: the Random Effect Probit estimator (REP).

In economics, Neural Networks have been principally used in two classes of applications: classification of economic agents and time series prediction[6]. Regarding classification, which is the subject of interest of this paper, ANN is widely diffused in bankruptcy prediction[7], while very few applications focus on financial crises. Nag and Mitra (1999) use a recurrent ANN, one for each country, to test its performance in predicting Malaysian, Thai, and Indonesian currency crises, and compare the result with those of signal approach. They find that the ANN model performs better than KLR. model, in particular when comparing out-of-sample prediction. Using data from East Asian countries, Franck and Schmied (2003) show that an optimized ANN does a better job in predicting currency crises with respect to logit, in particular signalling the currency crises that hit Russia and Brazil in the late nineties.

This paper develops a Feed-forward neural network for debt crises prediction and confronts its performances with those of a probit regression. The Universal Approximation Theorem guaranties, under particular condition, that it is possible to approximate, with arbitrary accuracy, the function that relates default probability and explanatory variables so that a well developed ANN should outperform others statistical techniques.

The organization of the paper is the following: Section 2 describes the dataset. Preliminary analysis and benchmarks selection (the random effect probit estimator) are developed in Section 3. Section 4 contains a brief

5 Various models use different definition of debt crisis, so that crises episodes could differ between authors.

6 There is a third, but less investigated, application that models bounded rational economic agent. See Herbrich et al. (1999) for the classification and a short review.

7 See for example Perez (2006).

description of ANNs' functioning, the performance of various models and a comparison in performances. Section 5 concludes.

2 DATA DESCRIPTION

The database consists of a wide set of information (34 row or transformed explanatory variables, with yearly frequency) regarding an (unbalanced) panel of 46 emerging countries in the period 1980-2004. The choice of variables is driven by the theory and the results of previous empirical works. Explanatory variables can be classified in internal (GDP growth, inflation, interest rate, etc.), external (US treasury bill interest rate, overvaluation, exchange rate agreement, degree of openness, etc.), and debt-related (average maturity, total external debt, short term external debt, interest on external debt, etc.). The effort is devoted to replicate the database used by Manasse et al. (2003), in the attempt to have a reference work to confront the results with. In particular, the dependent variable is defined in the same way as in the above cited paper:

> *"A country is defined to be in a debt crisis if it is classified as being in default by Standard & Poor's or if it receives a large nonconcessional IMF loan defined as access in excess of 100 percent of quota."*[8]

The default episodes are listed in Standard & Poor's (2004), while nonconcessional loan are drown from IMF International Financial Statistics (IFS) database. Explanatory variables are collected from IMF World Economic Outlook (WEO) and IFS databases, and World Bank Global Development Finance (GDF) database. Data on *de facto exchange rate agreements* are directly provided by the IMF's Public Affairs Division. Table 1 lists crisis episodes used to define the dependent variable, while Table 2 describes the explanatory variables used to predict those crises. The choice to use variables in relation to the GNI instead of GDP is aimed by two elements: the first is that countries' ability to repay theirs debt depends on their "own" income, while the second comes from preliminary regression results, in which variables in relation

[8] Manasse et al. (2003), pg. 8. The list of crisis episodes produced with the strict application of this rule is a bit different from that of Manasse et al., because they correct crisis episodes using their own information. Furthermore, some differences could be done by changes in the IMF and S&P's data.

to GNI appear to be more statistical significant than those calculate with respect to GDP. At the same time GDP appears to be a more significant regressor than GNI.

Tab. 1 **Crisis Episodies**

COUNTRIES	CRISIS EPISODIES
Algeria	1991-'00
Argentina	1982-'04
Bolivia	1980-'84, 1986-'99, 2003-'04
Brazil	1983-'94, 1998-'99, 2001-'04
Chile	1983-'91
China	
Colombia	
Costa Rica	1981-'90
Cyprus	
Czech Republic (*)	1993
Dominican Republic	1980-'01
Ecuador	1982-95, 1999-'00
Egypt	
El Salvador	1981-'96
Estonia (*)	1995-'96
Guatemala	1981-'84, 1986, 1989
Hungary (*)	1983-'87, 1991-'94
India	1982-'88, 1991-'93
Indonesia	1997-'04
Israel	
Jamaica	1980-'94
Jordan	1995-'04
Kazakhstan (*)	1995-'98
Latria (*)	1994-'95
Lithuania*	1994-'00
Malaysia	1982-'99
Mexico	1980-'91
Morocco	
Oman	
Pakistan	1980-'87, 1989-'04
Panama	1981-'96
Paraguay	1986-'92, 2003-'04
Peru	1980-'98
Philippines	1985-'94, 1997-'02
Poland (*)	1981-'94
Romania (*)	1981-'86, 1991, 1994
Russia (*)	1991-'00
Slovak Republic (*)	1993-'95
South Africa	1982, 1985-'87, 1989, 1993
Thailand	1980-'88, 1997-'01
Trinidad	1988-'91
Tunisia	1986-'89, 1991-'94
Turkey	1980-'87, 2000-'04
Ucraine (*)	1995-'02
Uruguay	1983-'88, 1990-'91, 2002-'04
Venezuela	1983-'88, 1990-'92,1995-'98

(*) For transition countries data on explanatory variables are not available prior to 1992.

Tab. 2 **Variables and Statistics**

	VARIABLE	OBS	MEAN	STD.DEV	MIN	MAX	LABEL
1	gdp_gr	1008	3.06	4.99	-32.10	17.10	GDP growth rate - (IMF - WEO)
2	gni_gr	959	5.84	14.27	-63.38	92.47	GNI growth rate - (WB GDF)
3	inf	1008	81.12	526.67	-30.30	11749.60	Inflation rate - (IMF -WEO)
4	mch	992	60.80	352.42	-40.08	6724.82	Money change, % - (IMF IFS)
5	ir	880	78.56	729.21	1.27	17235.80	Short term interest rate - (IMF IFS)
6	tbill	1104	6.22	3.03	1.01	14.08	US T-Bill interest rate - (IMF IFS)
7	aint	935	6.70	2.31	0.00	16.50	Average interest rate - (WB GDF)
8	overv	966	45.45	20.38	0.00	135.14	Overvaluation, % - (overv>100 = overvaluation)
9	exrag	626			1.00	8.00	De facto exchange rate agreement - (IMF)
10	amat	935	15.49	13.64	0.00	385.40	Average maturity, years - (WB GDF)
11	grace	935	5.09	2.44	-10.70	25.40	Average grace period, years - (WB GDF)
12	grant	935	18.97	14.30	-23.40	78.70	Average grant element, % - (WB GDF)
13	open	912	79.40	49.94	15.47	498.99	Degree of openess as % of GNI
14	tb_gni	912	-3.28	6.08	-26.20	24.61	Trade balance as % of GNI
15	def_gni	749	-3.09	5.13	-93.33	10.24	Deficit as % of GNI
16	ted_gni	920	53.93	32.23	0.14	253.22	Total external debt as % of GNI
17	ted_xgs	913	176.73	115.88	5.91	874.99	Total external debt as % of export
18	ted_res	931	694.91	809.08	29.46	8399.82	Total external deb as % of reserve
19	sted_gni	920	8.37	7.18	0.00	52.00	Short term external debt as % of GNI
20	sted_xgs	913	26.87	22.64	0.00	169.36	Short term external debt as % of export
21	sted_res	931	106.20	169.00	0.00	2399.83	Short term external debt as % of reserve
22	sted_ted	935	16.55	11.06	0.00	81.71	Short term external debt as % of total external debt
23	ited_xgs	913	10.03	8.13	0.02	55.69	Interest on ted as % of export
24	ited_res	931	38.96	50.25	0.66	627.42	Interest on ted as % of reserve
25	isted_xgs	913	1.72	2.35	0.00	30.42	Interest on sted as % of export
26	isted_res	931	6.20	11.60	0.00	190.37	Interest on sted as % of reserve
27	Tds_xgs	913	22.93	14.67	0.02	117.81	Total debt service as % of export
28	Tds_res	931	88.05	92.66	1.34	882.32	Total debt service as % of reserve
29	stds_xgs	913	3.13	3.35	0.00	30.92	Short term debt service as % of export
30	stds_res	931	13.14	23.17	0.00	209.12	Short term debt service as % of reserve
31	Cac_gni	911	-2.34	5.15	-23.24	18.52	Current accaunt as % of GNI
32	Cac_res	930	-34.74	92.39	-857.90	998.98	Current accaunt as % of reserve
33	Cac_pef	403	1554.92	73328.00	-830705.00	920940.00	Current accaunt as % of portfolio equity flow
34	Cac_fdi	903	206.74	18640.30	-60816.70	473420.00	Current accaunt as % of foreign direct investment

As can be seen, the dependent variable is unevenly distributed between crisis and non-crisis years[9] (Table 3), and shows a quite strong state-dependency (Table 4).

Tab. 3 **Distribution of the dependent variable**

State	Overall		Betweeen Countries	
	Frequency	Percentage	Frequency	Percentage
0	665	59.33	46	100.00
1	449	40.67	40	86.96
Total	1104	100.00		

Tab. 4 **Transition Probabilities**

t+1 / t		0	1	
	0	92.06	7.94	100.00
	1	12.70	87.30	100.00
	Total	58.98	41.02	100.00

The results of t-test (Table 5) show that chosen variables are good candidates to discriminate between crises and non-crises episodes; only few of these don't pass the group mean comparison test.

In the following analysis observations for independent variables at time *t* are used to predict the state of the dependent variable (crisis/non-crisis) at time *t+1*, so that the forecast is a one-step-ahead forecast. Hence the covariates in the final database range from 1980 to 2003, while the dependent variable ranges from 1981 to 2004.

9 Binomial test rejects the null of equal proportion between crises and non crises at all level of significance.

Tab. 5 **Mean Comparison Test**

Variable	Debt Crisis=0			Debt Crisis=1			Mean t-test	
	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.	t	prob
GDP growth rate	559	3.61	0.22	449	2.38	0.22	3.901	0.000
GNI growth rate	525	6.79	0.42	434	4.96	0.78	2.263	0.024
Inflation rate	559	63.28	23.50	449	103.34	23.03	-1.200	0.230
Money change	547	35.98	11.04	445	91.30	20.85	-2.465	0.014
Short term interest rate	489	18.02	1.46	391	154.28	55.09	-2.765	0.006
US T-Bill interest rate	654	6.08	0.13	450	6.41	0.27	-1.811	0.070
Average interest rate	497	6.52	0.11	438	6.90	0.10	-2.564	0.011
Overvaluation	550	49.35	0.94	416	40.29	0.81	7.012	0.000
Average maturity	497	15.19	0.80	438	15.82	0.28	-0.709	0.479
Average grace period	497	5.25	0.12	438	4.90	0.09	2.236	0.026
Average grant element	497	19.11	0.65	438	18.81	0.68	0.326	0.745
Degree of openess as % of GNI	489	83.47	2.17	423	74.70	2.52	2.653	0.008
Trade balance as % of GNI	489	-2.83	0.29	423	-3.80	0.27	2.429	0.015
Deficit as % of gni	407	-3.15	0.30	342	-3.00	0.21	-0.400	0.689
Total external debt as % of GNI	492	44.79	1.23	428	64.43	1.66	-9.668	0.000
Total external debt as % or res	486	132.67	3.67	427	226.88	6.23	-13.403	0.000
Total external deb as % of res	494	426.46	17.38	437	998.39	49.10	-11.499	0.000
Short term external debt as % of GNI	492	7.90	0.30	428	8.91	0.38	-2.137	0.033
Short term external debt as % of xgs	486	22.95	0.82	427	31.34	1.27	-5.677	0.000
Short term external debt as % of res	494	72.98	3.49	437	143.75	10.85	-6.517	0.000
Short term external debt as % of ted	497	18.69	0.54	438	14.12	0.44	6.441	0.000
Interest on ted as % of xgs	486	7.84	0.28	427	12.53	0.45	-9.095	0.000
Interest on ted as % of res	494	24.98	1.22	437	54.76	3.05	-9.441	0.000
Interest on sted as % of xgs	486	1.38	0.07	427	2.12	0.14	-4.792	0.000
Interest on sted as % of res	494	4.29	0.28	437	8.37	0.73	-5.438	0.000
Total debt service as % of xgs	486	20.27	0.60	427	25.95	0.75	-5.943	0.000
Total debt service as % of res	494	65.58	2.73	437	113.44	5.44	-8.136	0.000
Short term debt service as % of xgs	486	1.95	0.87	427	4.48	0.20	-12.254	0.000
Short term debt service as % of res	494	6.46	0.39	437	20.70	1.48	-9.826	0.000
Current accaunt as % of GNI	488	-2.30	0.24	423	-2.39	0.25	0.281	0.779
Current accaunt as % of res	492	-26.11	2.91	438	-44.42	5.51	3.027	0.003
Current accaunt as % of pef	247	154.42	2622.78	156	3772.39	8488.60	-0.482	0.630
Current accaunt as % of fdi	481	1230.30	1133.89	422	-959.94	295.27	1.764	0.078
De facto exchange rate agreement *	392			234			29.550	0.000

(*) Pearson chi^2 test.

3 PRELIMINARY ANALYSIS AND BENCHMARK SELECTION

The whole sample is used to investigate the statistical properties of the data, while in preliminary and successive estimations three subsample are exploited to scrutinize the forecasting performance (*hit rate*) of different techniques. The first subsample is the full sample. The estimation is based on all the observations, and those from 2002 to 2004 are used to assess the forecasting performance (in-sample forecast). In the second trial, parameters are estimated using the observation in the 1981-2001 interval, and the remaining are devoted to the hit rate evaluation (out-of-sample forecast). The last attempt uses the above out-of-sample forecasting interval, but parameters estimation is based on the 1990-2001 time spam. The use of three subsample permits to test whether there are strong differences between in-sample and out-of-sample forecast (first *vs* second subsample) and whether recent crises show different characteristics with respect to less recent ones (second *vs* third subsample), so that it is preferable to limit the estimates to the more recent information.

3.1 Selecting the benchmark

To assess the prediction aptitude of ANN it is necessary to compare its results with a benchmark. Various statistical and econometric techniques have been evaluated taking care on two main features: the probabilistic nature of data, and the prediction performance of the technique.

The nature of the data doesn't allow to use Linear Discriminant Analysis. In fact, covariates have to be independent, normally distributed, and with the same variance-covariance matrix between groups. Furthermore data must be linear separable. Preliminary statistics show that the collected covariates violate one or more of these hypothesis[10].

Data characteristics suggests to use a non-linear regression technique, and, in particular, those for binary response model. The comparison between pooled logit and the conditional (or Fixed Effect) logit shows that the former has a better performance in prediction, but the Hausman test indicates the presence of heterogeneity, inducing to adopt the latter. On the other hand, the weakness of the Chamberlain (1980) approach is that it discards a lot of information, since it uses only the pair of successive observation which lay in different state. One

---

[10] Preliminary statistics are omitted from the paper. They are available from the author upon request.

way to use the full set of available observation is to use the random effect probit (REP) estimator. Even thought the hypothesis of independence between unobserved heterogeneity and covariates seems to be very strong, in particular when dealing with countries, prediction performances suggest to use the REP estimator.

At the end of this phase some variables had been discarded (*de facto* exchange rate agreements, public deficit as percentage of GNI, and current account balance as a percentage of portfolio equity flows) because they have few observations and/or don't show significant differences in the mean.

3.2 The Random Effect Probit Estimator

The first step in the estimation of REP model, for all subsample, is to start with all variables, except the three above mentioned. Once the full model has been estimated, those variables that are not significant are stepwise deleted[11]. Columns 2, 4 and 6 (REP) of Table 6 show estimated coefficients for these unrestricted models. They are unrestricted because, even though all the coefficients are statistically significant, some coefficients show a counter intuitive sign in terms of economic influence. Odd columns of the same table report estimation results for the restricted models (Restricted REP), in which only those variables that have the correct sign in the REP model are maintained. The second part of Table 6 reports classification and prediction results[12]. Since the results of the probit are probabilities, a threshold to discriminate between crisis and non-crisis must be chosen. In these kind of models it is common to choose the relative frequency of crises (1s) in the estimation sample, which vary between different samples used.

Estimation results do not show a clear winner model, that is a common subset of variables which is always maintained. The only variables which are common and significant to all models are total external debt as a percentage of official reserves, and GDP growth. More importantly, the hit rate strongly deteriorates removing from the three REP models the variables with counter intuitive sign, and this phenomenon is particularly stronger when performing prediction rather than classification. The REP model estimated in the

[11] Manual stepwise procedure is necessary because the dataset is unbalanced. In fact, the automatic procedure for stepwise implemented in Stata9 uses in every step the smallest observation set common to the initial set of variables. Even though a common set of observation implies a direct comparability, the huge waste of information is very inefficient.

[12] Classification is referred to the observations used to estimate the model, while prediction is used for the period 2002-2004.

Tab. 6 Regression Results

	1981-2004		1981-2001		1990-2001	
	REP	Restricted REP	REP	Restricted REP	REP	Restricted REP
gdp_gr	-0.0665***	-0.0541***	-0.0450***	-0.0628***	-0.0840***	-0.0388**
mch	0.0021*	-0,0001			0.0025*	0,0004
ir	-0.0008**				-0.0010**	
tbill	0.1057***	0.0805***				
overv	-0.0122*				-0.0325**	
ted_gni	-0.0114*		-0.0105*		-0.0318***	
ted_xgs	0.0041**	0.0061***	0.0031**	0.0045***		
ted_res	0.0019***	0.0008***	0.0020***	0.0013***	0.0041***	0.0018***
sted_gni	0.0714**	-0,0143	0.0708**	-0,015		
sted_ted	-0.0563***		-0.0573***		-0.0758***	
isted_res	-0.0331***		-0.0321**			
tds_res	-0.0088***		-0.0064***			
stds_xgs	0.1836***	0,0464	0.1153***	0,0261		
cac_gni	0.0919***		0.1011***		0.3642***	
cac_res	-0.0060***	0,0004	-0.0080***	0,0009	-0.0142***	-0.0086**
cac_fdi	5.72e-06*		5.13e-06*			
amat			0.0246*			
open					0.0212***	
tb_gni					-0.1760**	0.1018***
sted_xgs					0.0537***	0,0106
ited_xgs					0.1190**	0,0156
ited_res					-0.0494***	
cons	-0,709	-2.0639***	-0.8988**	-1.3879***	-0,8631	-1.4566***
lnsig2u_cons	0,0576	-0,2005	0,168	0,0756	0.6517*	0,2934
rho	0,5144	0,45	0,5419	0,5189	0,6574	0,5728
sigma_u	1,0292	0,9046	1,0876	1,0385	1,3852	1,158
Obs.	755	891	755	768	444	473
pseudo_R^2(1)	0,3078	0,2295	0,2627	0,2007	0,2812	0,1602
loglikelihood	-296,7758	-388,981	-298,615	-334,5557	-173,174	-214,0605
chi^2_0(2)	263,9736	231,6641	212,7913	168,0139	135,5099	81,6528
chi^2_c(3)	104,7894	120,8966	138,5592	139,4646	77,5249	83,6651
AIC	629,5516	797,9619	627,23	685,1114	378,3479	446,1209
BIC	712,8325	845,8854	696,6308	722,2617	443,8811	483,5528
Obs. In Classification	755	891	755	768	444	473
Correctly Classified 0	81,82	80,75	80,60	79,32	80,54	78,55
Correctly Classified 1	72,09	70,10	73,07	65,28	72,73	66,16
Obs. In Prediction	118	127	127	127	118	127
Correctly predicted 0	92,63	94,12	90,20	92,60	77,89	87,25
Correctly predicted 1	60,87	52,00	68,00	52,00	78,26	48,00

* p ≤ 0.1; ** p ≤ 0.05; *** p ≤ 0.01.

(1) pseudo_R^2=((e(ll_0)-e(ll))/e(ll_0)).

(2) LR test: chi^2_0=-2*(ll-ll_0).

(3) LR test: chi^2_c=-2*(ll-ll_c).

where:

ll: loglikelihood of full REP model.

ll_0: loglikelihood of constant only REP model.

ll_c: loglikelihood of pooled probit model.

subsample 1990-2001 is the best performer because it reaches the higher hit rate in the forecasting horizon, correctly predicting 77.89% of non-crisis and 78.26% of crisis episodes in 2002-2004. It is important to note that, in evaluating the performance, type I error (predicting non-crisis in case of default) is a more serious mistake than type II error (predicting crisis in the case of a healthy nation). For this reason, it wold be sometimes recommended to minimize a loss function in which the two errors have different weight in order to select the best model. Since the aim of the paper is to confront REP with ANN, this further development is not considered.

4 ARTIFICIAL NEURAL NETWORK

The interest in artificial neural networks (ANNs), whose functioning resemble the logic of human brain, comes from the fact that it can be used in intensive data process to approximate function that are highly non linear. The learning process is based on the "experience" by which the ANN adjusts, using various kind of training processes, the weights that link together explanatory variables in order to get a particular goal. The elementary unit in ANN is the neuron, whose task is to process the row data, by means of appropriate weight and bias, and to apply a function[13], so that transformed data reproduce the target to the better. The training consists in the process of successive adjustments of weights and biases so that the distance, in terms of a specific error function, between the output of the neuron and the target is minimized. ANN with a single neuron and threshold activation function are called *perceptron*[14], while ANN with more than a single perceptron, arranged sequentially and in parallel, are called *Multi-layer perceptron* or *Back-propagation network*.

As previously mentioned, the training process can be thought as a sequence of successive adjustments of weights and biases in order to minimize the error function. Such a process involves two different steps: in the first the derivative of the error function with respect to weight and biases is computed. In this step the task of the Back-propagation algorithm is to get an efficient way to evaluate, using the appropriate backward chain rule, the derivative of error

[13] These functions are called activation functions.

[14] Rosenblatt (1962).

function with respect to weights and biases. In the second step the obtained derivatives are used to compute the adjustments needed to weights and biases. Various optimization algorithms are available for this task: gradient descent, conjugate gradient, quasi-Newton methods, and, most recently, Levenberg-Marquardt algorithm. They differ in the performance and in the efficiency of computation. Optimization algorithms are usually called training algorithm, even thought the training implies both first and second steps. Here the usual convention is followed, bearing in mind this difference.

In this work various tries, for different training algorithm inside the same basic architecture for a multi-layer perceptron, have been conducted[15]. The network is a two layers network[16] with 3 units in the hidden layer[17] and one unit in the output layer. Activation functions are log-sigmoidal (logistic); the performance function is the mean squared error (MSE); the network is trained for 100 epochs. The algorithm of the training process is Levemberg-Marquardt; Section 4.1clarifies the choice. The general formula for this network is:

$$y_k = g\left(\sum_{j=0}^{m} \omega_{kj}^{(2)} \cdot g\left(\sum_{i=0}^{n} \omega_{ji}^{(1)} x_i\right)\right) \tag{1}$$

where $g(a) = \dfrac{1}{1+\exp(-a)}$, ω_{ji} is the weight applied to input variable *i* of neuron *j* in the hidden unit, while ω_{kj} is the weight applied to the output of the *j-th* hidden unit before being processed by the *k-th* output unit – in this case unique.

The choice of the architecture is motivated by the fact that, in the context of classification problem, networks with logistic activation function, two layer of weight, and sufficient number of unit in the hidden layer, can approximate any decision boundary with arbitrary accuracy (*Universal Approximation Theorem*). Furthermore, this ANN architecture allows to model the output as posterior probability of class membership[18]. In choosing the number of epochs it is necessary to avoid the problem of overfitting. The higher the number of training epochs the better is the classification performance but, at the same time, the worse is the network ability to correctly predict the future processing data not

---

[15] This paper uses the same counting convention for the number of network layers used by Bishop (1995) pg. 119.

[16] The hidden layer and the output layer.

[17] Networks with more units in the Hidden layer are also tried, but they didn't show significant improvement in terms of hit rate. On the other hand, adding a neuron considerably lengthens the training process because it increases the number of weights and biases to be estimated.

[18] Bishop (1995), Sections 6.5 and 6.6.

seen before, i.e. networks generalization deteriorates. In order to get the optimal number of epochs, at a first step the above described network is trained for different number of epochs, setting weight and bias to zero. Once the optimal number of ephocs is obtained weights and biases are chosen randomly, as is usual in these models. This two steps procedure is necessary because, other things equal, the path of the network depends on initial condition, and, in order to confront the results for different training lengths.

Another way to alleviate the problem of overfitting is by using techniques able to improve the generalization of a network. The most diffused methods are *regularization* and *early stopping*. The application of automated Bayesian regularization, which imply a simple extension of the network, produces poor results. The implementation of early stopping is just a little bit more complicated. Given the nature of the data, a panel of 46 countries observed in the period 1981-2004, the observation for the validation set couldn't be chosen along just one dimension, but must be chosen randomly between time and countries. To allow for early stopping the following procedure has been applied.

In the first step dimensionality is reduced using as input variables those selected by the REP (column 4 Table 6); then observations are split in estimation sample (1981-2001) and forecasting sample (2002-2004); in the third step 100 observations are randomly drown from the estimation sample and used for validation, leaving the others for training. Also using this techniques the performance of the network does non improve[19].

4.1 Pre-Processing and Dimensionality Reduction

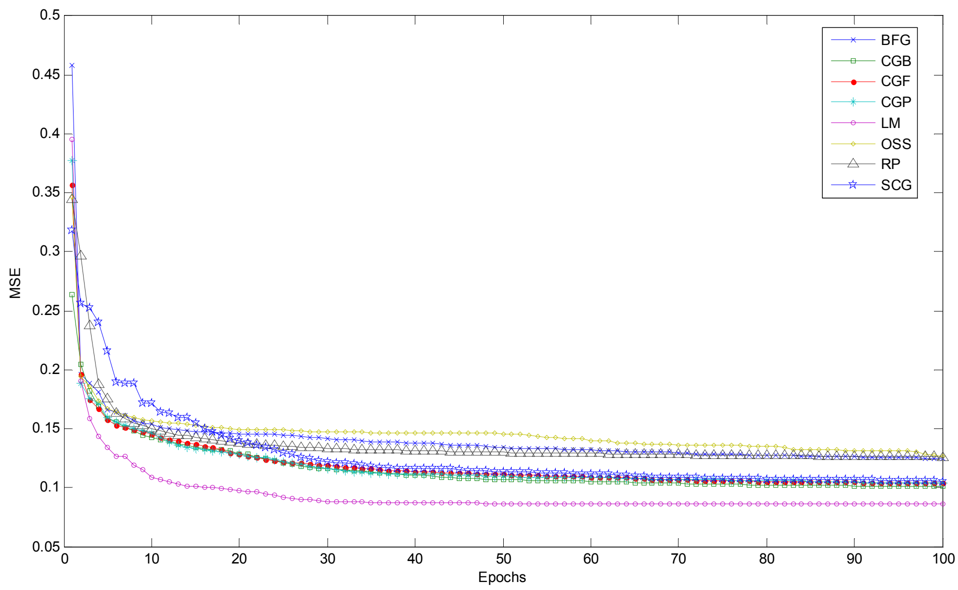
For most application it is suggested, if not even strictly necessary, to apply some sort of pre-processing to the whole dataset. Two of the most diffused forms of pre-processing are *dimensionality reduction* and *rescaling*. Dimensionality reduction alleviates the course of dimensionality because reduces the number of parameter to be estimated by discarding some input variables, and helps to improve generalization. Rescaling, instead, helps to eliminate the problem linked to the absolute value of different variables, which does not necessarily reflects their relative importance. Although in the dataset there are not variable with very different value, when dimensionality reduction

[19] When the number of observation is much greater than the number of parameters to be estimate, the problem of overfitting does not arise. So the lack of improvement in the performance of the network when generalization procedure are used.

technique is applied normalization is also used because the latter is strongly recommended when the former is used.

In order to evaluate the performance of different training algorithms, the normalization of the mean and standard deviation to zero and one respectively, to rescale the data, and principal component analysis (PCA), as a tool for reducing the curse of dimensionality, are applied. At this point the network is trained using different algorithms in order to test which is the best[20]. Figure 1 shows the speed of various algorithms in reducing the MSE.

Figure 1 **Algorithms' speed***



(*) Sample: full sample.
Preprocessing method: Principal Component Analysis.

Since the Levenberg-Marquardt seems to be the fastest and the most efficient algorithm, remaining analysis are developed using it. Once the best network architecture is defined[21], the network is trained and simulated for all subsample periods[22] and three different pre-processing methods: *i)* PCA; *ii)* REP; *iii)* Restricted REP. Since ANNs are sensible to initial conditions – i.e.

[20] The algorithms used are: Levenberg-Marquardt (LM); BFGS Quasi-Newton (BFG); Resilient Backpropagation (RP); Scaled Conjugate Gradient (SCG); Conjugate Gradient with Powell/Beale Restarts (CGB); Fletcher-Powell Conjugate Gradient (CGF); Polak-Ribiére Conjugate Gradient (CGP); One-Step Secant (OSS).

[21] Table 7 resumes the main characteristics of the network.

[22] Those used for REP and restricted REP.

starting weight and bias – and the aim of the paper is to show that, under certain conditions, an ANN can outperforms the probit, a symple expedient is used: the network is nested in a *while* loop so that the loop stops if and only if classification and prediction performance of the network are at least as good as those of the corresponding probit. The scheme of the loop is the following:

1. Set the weights and biases randomly;
2. Train the network until epoch 100;
3. Tabulate classification and prediction results in terms of crisis (1) and non-crisis (0) using the proper threshold;
4. If classification and prediction results are at least as good as those of the probit end the loop, otherwise go to step 1 and try again.

The Universal approximation theorem guarantees that this loop is not an infinite loop[23].

Tab. 7 Elements of the architecture of the Artificial Neural netvork

1	Type of network: **Multi-layer perceptron or Back-propagation network**
2	N. of Layers: **2**
3	Number of neurons in the hidden layer: **3**
4	Number of neurons in the output layer: **1**
5	Activation functions: **Logistic**
6	Performance function: **Mean Squared Error**
7	Training algorithm: **Levemberg-Marquardt**
8	Starting weights and biases: **Random**
9	N. of training ephocs; **100**

[23] This is a very strong sentence indeed. In fact the Universal Approximation Theorem state that in the context of classification problem, networks with logistic activation function and two layer of weight and sufficient number of unit in the hidden layer, can approximate any decision boundary with arbitrary accuracy, but it is not possible a priori to know which is the sufficient number of unit in the hidden layer. At least in this case 3 seams to be the right number! Furthermore not all algorithms listed in note 12 outperform the probit in a finite number of loops, but it can be a problem related to the architecture of the network. On the other hand the Levemberg-Marquardt algorithm outperforms the probit, in various experiments in less than 10 loops.

The thresholds for the networks that use, as a pre-processing method, variables selected by REP and Restricted REP are those of the corresponding probit, while the network which uses the PCA has the same thresholds of the REP, which are generally higher with respect to those of the Restricted REP.

4.2 Forecasting Results

Table 8 shows network forecasting performance for various pre-processing methods across different training periods. Given the expedient used in stopping the process, each one of these outperform the correspondent probit.

Tab. 8 ANNs' Classification and Prediction Results using Levemberg-Marquardt algorithm

Pre-processing method			1981-2004	1981-2001	1990-2001
Principal component analysis	Classification	Correctly Classified 0	92.77	92.22	87.45
		Correctly Classified 1	85.23	85.43	91.80
	Prediction	Correctly Predicted 0	95.79	90.53	80.00
		Correctly Predicted 1	91.30	69.57	86.96
REP	Classification	Correctly Classified 0	84.15	83.68	83.23
		Correctly Classified 1	81.60	80.80	77.54
	Prediction	Correctly Predicted 0	92.63	91.18	80.00
		Correctly Predicted 1	82.61	80.00	78.26
Restricted REP	Classification	Correctly Classified 0	81.99	80.37	92.00
		Correctly Classified 1	75.74	74.34	74.75
	Prediction	Correctly Predicted 0	95.10	93.14	87.25
		Correctly Predicted 1	68.00	56.00	52.00

Between different pre-processing methods the one which involves normalization and PCA seems to be the most effective. Obviously the hit rates obtained using the full sample are higher because the network uses in-sample information in the prediction horizon. It can be seen that in-sample prediction performance of the network which uses the PCA reaches exceptionally high results: it correctly predict 95.79% of non-crises and 91.30% of crises episodes. In the out-of-sample prediction the network which uses the 1990-2001 sample for the training is slightly better the other one because it has a lower type I error: this result suggests that more recent crises have different features respect to those of the full sample, and this result is supported by that of the REP regression.

Between the remaining two subsample it is not clear which one is preferable because different performances depend on pre-processing methods, initial condition and stopping criteria in the *while* loop. The network which uses REP as a pre-selection shows good performance in out-of-sample prediction, correctly predicting 80% of crisis and emitting only 8.82% of false signals. The network which uses restricted REP reaches poor results, but this could be the consequence of fixing a very low threshold, equal to the performance of the corresponding probit.

The hit rate of the network can be improved increasing the thresholds in the *while* loop. For example, using the network with PCA, the period 1981-2001 for training and the remaining for validation, setting classification thresholds equal to the hit rate of REP for the same period, and increasing prediction thresholds to 85% both for correct prediction of crisis and non-crisis gives the following results: correctly classified non-crises: 92.51%; correctly classified crises: 88.41%; correctly predicted non-crises: 85.26%; correctly predicted crises: 91.30%.

A further comparison can be done with the results obtained by Manasse et al (2003). In their work they find that conjoint use of CART and probit permits to correctly *classify* in the full sample 81% of crisis episodes sending only 6% of false alarms. Setting these thresholds in a network with PCA and using the full sample permits to correctly *classify* 85.23% of crises and 94.17 of non crisis episodes.

5 CONCLUSION

This paper shows that, thanks to the Universal Approximation theorem, an ANN with only two layers can outperform a traditional Early Warning System in predicting a sovereign debt crisis, if one chooses the right number of hidden unit, training epochs and an efficient training algorithm. This is possible because if the relation that links debt crisis indicator and explanatory variables is highly non linear, the flexibility of ANNs should give, theoretically, results that are at least as good as those of the parametric traditional methods. Further refinements can be done to improve the ANNs' performance, for example using a committee of network. The characteristic of a committee is that it does perform at least as well as the best single network.

The drawback of ANNs is that they don't offer an immediate intuition for policy implication. In fact it is not straightforward to interpret, for example, the marginal effect, in terms of change in crisis probability, of an increase in an independent variable. It doesn't mean that it is always impossible, but the high non-linearity does not guarantee an explicit solution. On the other hand, ANNs do not pretend to be policy models, but just forecasting models.

Neither public nor private institutions entrust their prediction in a single automatic procedure. For example the IMF use a dual-core early warning system which contains both a parametric (probit) and a non parametric (signal) model[24]. The flexible functional form of ANNs can improve the performance of an early warning system because of its strong data processing ability, but, at the same time, should be used together with traditional methods and, last but not least, together with "real" human brain.

[24] IMF (2002).

REFERENCES

Abiad, A. (2003) “Early-Warning Systems: A Survey and a Regime-Switching Approach”, IMF Working Paper N. 32.

Berg, A., Borensztein, E. and Pattillo, C. (2004) “Assessing Early Warning Systems: How They Have Worked in Practice”, IMF Working Paper N. 52.

Bishop, C.M., (1995) Neural Network for Pattern Recognition, Oxford University Press.

Chamberlain G. (1980) “Analysis of Covariance with Qualitative Data” Reviews of Economic Studies, Vol. 47 N. 1, pp 225-238.

Ciarlone, A. and Trebeschi, G (2004) “Currency and Debt Crises: A Review of the Early Warning Systems”. In: Wilkin, S. (Ed.), Country and Political Risk: Practical Insight for Global Finance (Risk Book).

Ciarlone, A. and Trebeschi, G (2005) “ Designing an Early Warning System for Debt Crises”, Emerging Maket Review, Vol. 6 N. 4.

Detragiache, E. and Spilimbergo, A. (2001) “Crises and Liquidity: Causes and Interpretation” IMF Warking Paper N. 2.

Franck, R. and Schmied, A. (2003) “Predicting Currency Crisis Contagion from East Asia to Russia and Brazil: An Artificial Neural Network Approach”, Bar-Ilan University, AMCB Working Papers N.2.

Frankel, J.A. and Rose, A.K. (1996) “Currency Crashes in Emerging Markets: An Empirical Treatment”, Journal of International Economics, Vol. 41, pp 351-366.

Herbrich, R., Keilbach, M., Graepel, T., Bollmann-Sdorra, P. and Obermayer, K. (1999) “Neural Networks in Economics”, in Advanced in Computational Economics: Computational Techniques for Modelling Learning in Economics, Kluwer Academics.

Kaminsky, G.L. (2003) “Varieties of Currency Crises” NBER warking Paper N. 10193.

Kaminsky, G.L, Lizondo, S. and Reinhart, C.M. (1998), “Leading Indicators of Currency Crises”, IMF Staff Papers, Vol. 45 N. 1.

Krugman, P. (1999) “Balance Sheets, the Transfer Problem, and Financial Crises”, mimeo.

Krugman, P. (2001) “Crises: The Next Generation”, mimeo.

IMF (2002) “Early Warning System Models: The Next Steps Forward”, Global Financial Stability Report, March.

Manasse, P., Rubini, N. and Schimmelpfenning, A. (2003) “Predicting Sovereign Debt Crises”, IMF Working Paper N. 221.

Nag, A.K. and Mitra, A. (1999) “Neural networks and early warning indicators of currency crisis”, Reserve Bank of India Occasional Papers, Vol. 20 (2), pp. 183-222.

Perez, M. (2006) “Artificial neural networks and bankruptcy forecasting: a state of the art”, Neural Computing & Application, Vol. 15 N. 2, pp 154-163.

Rosenblatt (1962) Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, Washington DC: Spartan.

Roubini, N. and Setser, B. (2005) “Bailouts or Bail-Ins?: Responding to Financial Crises in Emerging Economies”, Ed. Institute for International Economics.

Standard & Poor’s (2004) “Sovereign Defaults Set to Fall Again in 2005”. RatingsDirect, September 28.

Sturzenegger, F. and Zettelmeyer, J. (2006) Debt Defaults and Lessons from a Decade of Crises, Cambridge, Mass: MIT Press.

Working Papers available:		
n. 31/03	S. DE NARDIS C. VICARELLI	The Impact of Euro on Trade: the (Early) Effect Is not So Large
n. 32/03	S. LEPROUX	L'inchiesta ISAE-UE presso le imprese del commercio al minuto tradizionale e della grande distribuzione: la revisione dell'impianto metodologico
n. 33/03	G. BRUNO C. LUPI	Forecasting Euro-area Industrial Production Using (Mostly)\ Business Surveys Data
n. 34/03	C. DE LUCIA	Wage Setters, Central Bank Conservatism and Economic Performance
n. 35/03	E. D'ELIA B. M. MARTELLI	Estimation of Households Income from Bracketed Income Survey Data
n. 36/03	G. PRINCIPE	Soglie dimensionali e regolazione del rapporto di lavoro in Italia
n. 37/03	M. BOVI	A Nonparametric Analysis of the International Business Cycles
n. 38/03	S. DE NARDIS M. MANCINI C. PAPPALARDO	Regolazione del mercato del lavoro e crescita dimensionale delle imprese: una verifica sull'effetto soglia dei 15 dipendenti
n. 39/03	C. MILANA ALESSANDRO ZELI	Productivity Slowdown and the Role of the Ict in Italy: a Firm-level Analysis
n. 40/04	R. BASILE S. DE NARDIS	Non linearità e dinamica della dimensione d'impresa in Italia
n. 41/04	G. BRUNO E. OTRANTO	Dating the Italian Business Cycle: a Comparison of Procedures
n. 42/04	C. PAPPALARDO G. PIRAS	Vector-auto-regression Approach to Forecast Italian Imports
n. 43/04	R. DE SANTIS	Has Trade Structure Any Importance in the Transmission of Currency Shocks? An Empirical Application for Central and Eastern European Acceding Countries to EU
n. 44/04	L. DE BENEDICTIS C. VICARELLI	Trade Potentials in Gravity Panel Data Models

Working Papers available:		
n. 45/04	S. DE NARDIS C. PENSA	How Intense Is Competition in International Markets of Traditional Goods? The Case of Italian Exporters
n. 46/04	M. BOVI	The Dark, and Independent, Side of Italy
n. 47/05	M. MALGARINI P. MARGANI B.M. MARTELLI	Re-engineering the ISAE manufacturing survey
n. 48/05	R. BASILE A. GIUNTA	Things change. Foreign market penetration and firms' behaviour in industrial districts: an empirical analysis
n. 49/05	C. CICCONI	Building smooth indicators nearly free of end-of-sample revisions
n. 50/05	T. CESARONI M. MALGARINI G. ROCCHETTI	L'inchiesta ISAE sugli investimenti delle imprese manifatturiere ed estrattive: aspetti metodologici e risultati
n. 51/05	G. ARBIA G. PIRAS	Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects
n. 52/05	L. DE BENEDICTIS R. DE SANTIS C. VICARELLI	Hub-and-Spoke or else? Free trade agreements in the "enlarged" European Union
n. 53/05	R. BASILE M. COSTANTINI S. DESTEFANIS	Unit root and cointegration tests for cross-sectionally correlated panels. Estimating regional production functions
n. 54/05	C. DE LUCIA M. MEACCI	Does job security matter for consumption? An analysis on Italian microdata
n. 55/05	G. ARBIA R. BASILE G. PIRAS	Using Spatial Panel Data in Modelling Regional Growth and Convergence
n. 56/05	E. D'ELIA	Using the results of qualitative surveys in quantitative analysis
n. 57/05	D. ANTONUCCI A. GIRARDI	Structural changes and deviations from the PPP within the Euro Area

Working Papers available:		
n. 58/05	M. MALGARINI P. MARGANI	Psychology, consumer sentiment and household expenditures: a disaggregated analysis
n. 59/05	P. MARGANI R. RICCIUTI	Equivalenza Ricardiana in economia aperta: un'analisi dinamica su dati panel
n. 60/05	M. BOSCHI A. GIRARDI	Euro Area inflation: long-run determinants and short-run dynamics
n. 61/05	M. BOVI	Book-Tax Gap. An Income Horse Race
n. 62/06	M. BOVI	The Cyclical Behavior of Shadow and Regular Employment
n. 63/06	G. BRUNO C. LUPI C. PAPPALARDO G. PIRAS	The cross-country effects of EU holidays on domestic GDP's
n. 64/06	M. COZZOLINO F. DI NICOLA M. RAITANO	Il futuro dei fondi pensione: opportunità e scelte sulla destinazione del TFR
n. 65/06	S. LEPROUX M. MALGARINI	Clima di fiducia e spesa delle famiglie in Italia: un'analisi disaggregata secondo il reddito degli intervistati
n. 66/06	M. BOVI	Consumers Sentiment and Cognitive Macroeconometrics Paradoxes and Explanations
n. 67/06	G. ROCCHETTI	Modelli di business nel mercato del software e partecipazione delle imprese italiane al fenomeno open source
n. 68/06	L. CROSILLA	La stagionalità delle inchieste ISAE su imprese e consumatori: aspetti metodologici ed evidenza empirica
n. 69/06	C. OLDANI	Money demand & futures
n. 70/06	R. BASILE S. DE NARDIS A. GIRARDI	Pricing to market of italian exporting firms
n. 71/06	B.M. MARTELLI G. ROCCHETTII	The ISAE Market Services Survey: Methodological Upgrading, Survey Reliability