Early Warning Systems: an approach via Self Organizing Maps with applications to emergent markets

Marina RESTA [a,1],
[a] *University of Genova, Italy.*

Abstract. Over the past two decades the globalization of market economies have led to a large number of financial crises in emerging markets. The case of Paraguay in earlier '90 of the past century or, more recently, the crises in Turkey, Argentina, and Far East Asian markets have taught the important lesson that such phenomena, originally arising at local basis, can spread contagiously to other markets as well. At the same time, this made clear the importance of Early Warning System (EWS) models to identify economic and financial vulnerabilities among emerging markets, and, ultimately, to anticipate such events. With this in mind, we have introduced an EWS model based on the powerful clustering capabilities of Kohonen's Self Organizing Maps. Using macroeconomic data of several emerging countries, our analysis has been twofold. We have originally provided a static snapshot of countries in our dataset, according to the way their macroeconomic data cluster in the map. In this way, we were able to capture the (eventual) reciprocities and similarities from various emerging markets. As second step, we have dynamically monitored their evolution path in the map over the time. As main results, we were able to develop a crisis indicator to measure the vulnerability of countries, and we have also provided a proper framework to deduce probabilities of future crises.

Keywords. Early Warning Systems, Self Organizing Maps, Emergent markets

1. Introduction

Over the past two decades a large number of emerging market economies was interested by financial crises whose economic, social and political consequences have been often devastating, especially because instead of remaining bounded to the country where originated, they widely spread to neighbour economies. In order to provide a brief review, moving towards a chronological order, we can start with the crisis of the European Monetary System (EMS) in 1992/1993. In 1994 Latin America knew the so called tequila effect, when the crisis hitting the Mexican economy spred to the neighbouring countries. In the middle of 1997, it was then the turn of many Asian economies (Thailand, Malaysia, Indonesia, the Philippines, Hong Kong, Singapore, Taiwan, and South Korea) that were interested by imbalances originated by the devaluation of the Thai baht. In 1998 the Rus-

[1]Marina Resta, DI.E.M. sezione di Matematica Finanziaria, via Vivaldi 5, 16126 Genova, Italy; E-mail: mresta@unige.it.

sian crisis was really spectacular too: it was the first time that a literature appeared about the bankruptcy of a state. At the dawn of the 2000's financial instability interested Brazil (1999), Argentina (2001, 2003), Turkey (2001), and South Korea (2005).

Early Warning Systems (EWS) emerged then as new tools for policy makers and for regulators, to anticipate whether and when some country or economic sector may suffer a financial crisis, or, more generally, to detect intrinsec economic weaknesses and vulnerabilities. In their general acception, EWS are nothing but econometric models, including both macro and microeconomic variables. Important issues are therefore related to the choice of data to evaluate, and of the underlying methodology to transform financial and macroeconomic indicators. However, since the critical episodes often exhibit various characteristics, the generalization through an economic model to explain all these different cases has been commonly assumed to be not possible, and the remedies have been distinguished according to the sources of the crisis. In particular, the attention of researchers has devoted to the development of EWS focused on the study of the financial contagion channel. Jagtiani et al. [5], for instance, developed an EWS model that provides signals of capital inadequacy in banks; using a range of financial indicators, [6] extracted early signals as well. Discrete choice models with continuous variables on the right-hand side (logit and probit models) have been also popular [1].

Within the framework depicted above, our work is intended to add valuable contributions towards different directions. Firstly, we have analised the situation of a panel of countries under a huge number of indicators, not only of financial type, but also economic, monitoring both internal (local) and external fiscal conditions whose imbalance may compromise the economic and financial stability of a country and its surroundings. In practice, this means to control various type of vulnerability at one time. The tool employed is a non parametric method based on unsupervised neural networks, namely: Self Organizing Maps (SOM). SOMs appeared to be a "natural" tool of investigation to our purposes, since they basically realize a non–linear projection of multi–dimensional data into a bi–dimensional lattice of nodes arranged according to their similarity. This gives us the possibility to exploit the degree of vulnerability of each observed country, and its capability to spread contagiously to neighbour countries. Here the notion of neighbourhood has been assumed into a wider acception, since it can be intended either in a geographical sense, or in the sense to incorporate countries sharing similar features from the economical and financial perspective.

The paper proceeds as follows. Section 2 starts by reviewing Self Organizing Maps. Section 3 after a brief description of the data sample, discusses the results obtained by our methodology. Section 4 then concludes.

2. Some brief remarks on Self Organizing Maps

The Self Organizing Map (SOM) [9] is a projection method based on the principle of space representation through dimension reduction: a finite set of input patterns is represented by means of a number of nodes (neurons), sharing with inputs the same format, and arranged into a mono or bi–dimensional grid; in order to avoid hedges effects, wraparound versions can be also implemented [8,10]. When an arbitrary input is presented to a SOM, a competitive procedure starts, during which a winner or leader neuron is chosen in the map, as the best matching node, according to a metric previously fixed.

A generic step of the procedure may be then summarized as follows: we will refer to the case of a mono–dimensional SOM, but the layout presented can be easily generalized to higher dimensional grids. If $\mathbf{x}(t) = \{x_j(t)\}_{j=1,\ldots,n} \in \mathbb{R}^n$ is the input item presented to a map M with q nodes with weights $\mathbf{m}_i(t) = \{m_{i,j}(t)\}_{j=1,\ldots,n} \in \mathbb{R}^n$, $(i = 1, ..., q)$, i^*_t will be claimed the winner neuron at step t iff:

$$i^*_t = arg\min_{i\in M}\left(\sum_{i\in M}\sum_{j=1}^{n}|x_j(t) - m_{i,j}(t)|^p\right)^{1/p} \quad , \; p \in \mathbb{N} \tag{1}$$

Note that although $p \in \mathbb{N}$, most common choices are for $p = 1$, or $p = 2$. Once the leader has been identified according to Eq. (1), the correction of nodes in the map takes place; if $N_{i^*}(t)$ is the set of neurons in the map belonging to the *neighbourhood* of i^* (in a topological sense), then:

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{i*,i}(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] \tag{2}$$

Here $h_{i^*,i}(\cdot)$ is an interaction function, governing the way the nodes adjust respect to the winning neuron on the grid. Typical shapes for h include the constant function:

$$h_{i^*,i}(t) = \begin{cases} \alpha,\, i = {i^*}_t \vee i \in N_{i^*}(t) \\ \\ 0,\ \text{otherwise} \end{cases} \tag{3}$$

with $\alpha \in (0, 1)$, and the Gaussian function:

$$h_{i^*,i}(t) = exp\left\{ -\frac{\displaystyle\sum_{r=1}^{n}|m_{i,r}(t) - m_{i^*,r}(t)|^2}{2} \right\},\ i = 1, \ldots, q \tag{4}$$

After iterating the procedure over a number of epochs, the map should tend to a steady organized state [3], and neighbouring neurons should represent similar inputs.

3. Case study

3.1. Description of the input set

Our dataset is a $51 \times 49 \times 51$ tensor made up by 51 countries observed through 49 different variables, from the first quarter of 1995 to the third quarter of 2007 (51 records). Table 1 lists the countries included in our panel.

Table 1.Countries belonging to our panel data. The table shows the countries included in our study, the corresponding shortname used, and the dataset they belong to.

Country	ID	Dataset	Country	ID	Dataset	Country	ID	Dataset
Argentina	ARG	WLD1, WLD2	Albania	ALB	WLD1, WLD3	China	CHI	WLD1,WLD4
Brazil	BRA	WLD1, WLD2	Bosnia Herzegovina	BOH	WLD1, WLD3	Hong Kong	HOK	WLD1, WLD4
Chile	CHI	WLD1, WLD2	Croatia	CRO	WLD1, WLD3	South Korea	SKO	WLD1, WLD4
Colombia	COL	WLD1, WLD2	Macedonia	MAC	WLD1, WLD3	Taiwan	TAI	WLD1,WLD4
Paraguay	PAR	WLD1, WLD2	Serbia	SER	WLD1, WLD3	Indonesia	IND	WLD1,WLD4
Peru	PER	WLD1, WLD2	Russia	RU	WLD1, WLD3	Malaysia	MAL	WLD1,WLD4
Uruguay	URU	WLD1, WLD2	Ukraine	UKR	WLD1, WLD3	Philippines	PHI	WLD1,WLD4
Venezuela	VEN	WLD1, WLD2	Armenia	ARM	WLD1, WLD3,WLD4	Thailand	THA	WLD1,WLD4
Bulgaria	BUL	WLD1, WLD3	Azerbaijan	AZE	WLD1, WLD3,WLD4	Vietnam	VIE	WLD1,WLD4
Estonia	EST	WLD1, WLD3	Belarus	BEL	WLD1, WLD3	Singapore	SIN	WLD1,WLD4
Hungary	HUN	WLD1, WLD3	Georgia	GEO	WLD1, WLD3,WLD4	India	IDI	WLD1,WLD4
Poland	POL	WLD1, WLD3	Kazakhstan	KAZ	WLD1, WLD3,WLD4	Pakistan	PAK	WLD1,WLD4
Latvia	LAT	WLD1, WLD3	Kyrgyzstan	KYR	WLD1, WLD3,WLD4	Iran	IRN	WLD1,WLD4
Lithuania	LIT	WLD1, WLD3	Moldova	MOL	WLD1, WLD3	Turkey	TUR	WLD1,WLD4
Romania	ROM	WLD1, WLD3	Tajikistan	TAJ	WLD1, WLD3,WLD4	Egypt	EGY	WLD1, WLD4
Slovakia	SLK	WLD1, WLD3	Turkmenistan	TUK	WLD1, WLD3,WLD4	Tunisia	TUN	WLD1, WLD4
Slovenia	SLO	WLD1, WLD3	Uzbekistan	UZB	WLD1, WLD3,WLD4	Kenya	KEN	WLD1, WLD4

For each country we have reported the corresponding label used in our analysis, and the subgroup for which the country has been taken into account. This point represents one of the key issues of this study. Our aim was in fact two–fold. We were interested to monitor whether or not financial/economic distress can spread over neighbour countries, or even to world scale. To such purpose, we analysed our sample data as a whole (WLD1), but we have also studied subgroups, obtained once WLD1 was split following (straightforward) territorial or economical divisions: this led us to consider Latin American countries (WLD2), Euro–area surrounding countries (WLD3), Asian (or Asian influenced) countries (WLD4).

For what it concerns the variables in use, most of our dataset was extracted from the World Bank database, although some data were taken from the International Monetary Fund (IMF), and from the Organisation for Economic Co–operation and Development (OECD) database. The sample was selected (with respect to choice of both country and time) to maximize data availability; however, some observations are missing for individual variables. Despite from standard approaches to EWS building that focus only on the determinants of a particular crisis type (financial, polytical, economical), our analysis takes into account factors that variously affect the general equilibrium of a country. The variables we have considered belong to four categories:

(i) foreign variables, like interest rates, nominal rate of exchange (with US Dollar assumed as currency benchmark), and foreign direct investment (FDI);
(ii) domestic macroeconomic indicators, such as monetary and fiscal shocks indicators, real Gross Domestic Product (GDP) growth rate, fiscal stance, public debt (% GDP), inflation rate, domestic investment ratios, and M1, M2 indicators;
(iii) external variables, such as over-valuation, the level of indebtedness, fixed–rate vs. floating-rate borrowing, and domestic–currency vs foreign–currency denomination;
(iv) debt components, divided into the amounts lent by commercial banks, confessional, variable–rate; short–term; and lent by multilateral developmentbanks (including the World Bank and regional banks, but not the IMF).

3.2. Discussion of the results

We have originally provided a static snapshot of countries in our dataset, according to the way the corresponding macroeconomic data cluster in the map; to such purpose, WLD1 and WLD2 were trained on hexagonal 35×9 maps, WLD3 and WLD4 on 24×9 maps. The map dimensions vary according to the subgroup of data in use, and they were chosen after a preliminary data snooping procedure, during which we examined the sensitivity of convergence indexes to changes in the number of neurons [2,4]. Figure 1 reports the results obtained on the four groups of data with variables referring to the third quarter of 2007.

Figure 1 is organized with two maps placed on each row: the first map is the classical Uniformity Matrix (U–Matrix), where we have put into evidence the relative importance of each node, whose size have been scaled by the distance matrix calculated from all 49 components. The map on the right hand side is the same U–Matrix where each node shows the countries that belong to it. Looking at the organization of the maps, one can observe that the critierion joined by the SOM algorithm is essentially a geographical one: see for instance Figure 1(a), where Armenia, Moldova and Iran are grouped into

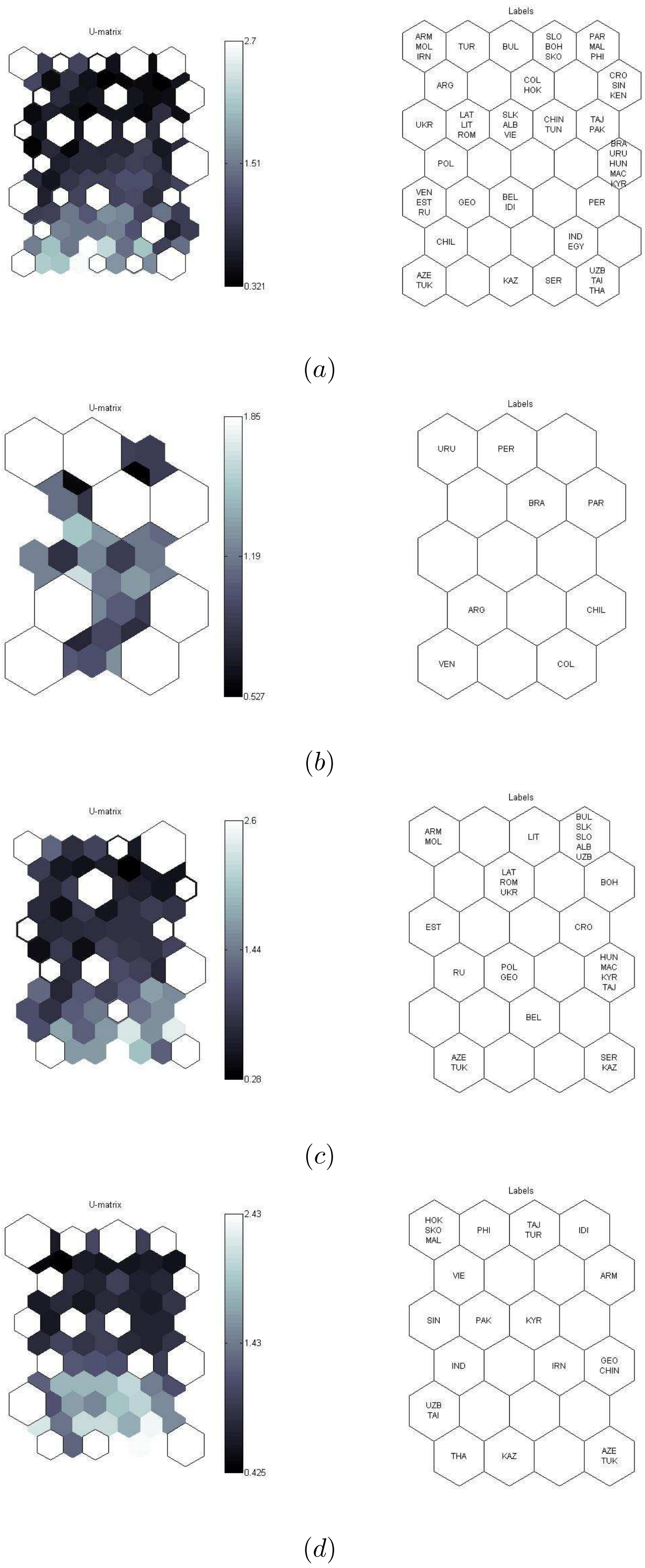


Figure 1. Results for WLD1 (a), and WLD2 (b), WLD3 (c) and WLD4 (d) datasets

Table 2. Probability of crisis in the countries belonging to our sample 8 quarters ahead.

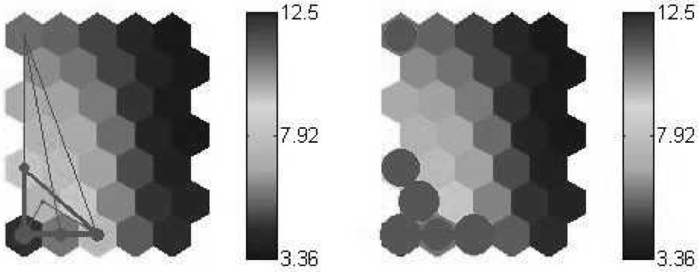
Country	ID	Prob	Country	ID	Prob	Country	ID	Prob
Argentina	ARG	40%	Albania	ALB	55%	China	CHI	40%
Brazil	BRA	45%	Bosnia Herzegovina	BOH	57%	Hong Kong	HOK	55%
Chile	CHI	30%	Croatia	CRO	45%	South Korea	SKO	50%
Colombia	COL	30%	Macedonia	MAC	55%	Taiwan	TAI	50%
Paraguay	PAR	30%	Serbia	SER	55%	Indonesia	IND	55%
Peru	PER	35%	Russia	RU	68%	Malaysia	MAL	65%
Uruguay	URU	35%	Ukraine	UKR	65%	Philippines	PHI	45%
Venezuela	VEN	45%	Armenia	ARM	56%	Thailand	THA	45%
Bulgaria	BUL	20%	Azerbaijan	AZE	60%	Vietnam	VIE	40%
Estonia	EST	20%	Belarus	BEL	60%	Singapore	SIN	48%
Hungary	HUN	20%	Georgia	GEO	61%	India	IDI	40%
Poland	POL	20%	Kazakhstan	KAZ	64%	Pakistan	PAK	70%
Latvia	LAT	20%	Kyrgyzstan	KYR	64%	Iran	IRN	48%
Lithuania	LIT	20%	Moldova	MOL	60%	Turkey	TUR	45%
Romania	ROM	20%	Tajikistan	TAJ	60%	Egypt	EGY	30%
Slovakia	SLK	20%	Turkmenistan	TUK	66%	Tunisia	TUN	30%
Slovenia	SLO	20%	Uzbekistan	UZB	66%	Kenya	KEN	55%

the same cell, and they are surrounded by Turkey (on the right) and by Turkmenistan and Azerbaijan (on the upper bound). This observation holds also for the other maps in Figure 1. Moving to the analysis of nodes, one should note that the maps sub (a) are those with the largest degree of generality, since clusters have been obtained using all the countries in the data sample: we therefore refer our conclusions mostly to such maps. According to the snapshot taken at the third quarter of 2007, nodes in the lower right side of the map refer to countries with reduced default probability. Such probability tends to increase moving to the upper side of the map, from left to right: the outcomes we reported on are somewhat confirmed by looking at the relative position of the countries in the maps sub (b)–(d).

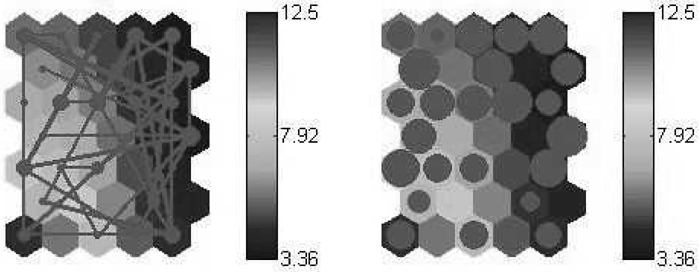
As second step, we have dynamically monitored the evolution path in the map of each country over the available 51 quarters, according to the different levels of significant variables. In this case, we traced the route on the map of each country (through the values of the 49 variables representing the related vector code), and we have inferred the probability of future crises 8 quarters (2 years) ahead. Figure 2 gives an idea of how the procedure works: on the upper side the dynamic route followed by Brazil is traced from 1995 to 2007; the lower part of Figure 2 shows the path of China within the same period. The probabilistic palette to asses the level of vulnerability of various countries in our sample is provided in Table 2.

4. Conclusions

This paper discussed a new Early Warning System (EWS) model, based on Self Organizing Maps (SOMs). Respect to existing EWS models, our model includes both economic and financial variables, and it is therefore better suited to provide signals of alarm respect to a wider range of imbalances and weaknesses a country may be exposed to. Addition-



(a)



(b)

Figure 2. Sample trajectories of Brazil (a) and China (b): the maps in the left hand side show details of the route traced through 51 quarters, the maps in the right hand side report the significance of node joined through the path.

ally, the method we have employed is non parametric, and it performs a non–linear projection of high dimensional input datasets into a bi–dimensional manifolds, thus being more robust than traditional logit or probit models that need preliminary data cleanings.

Our model was tested on data of 51 countries, mostly extracted from the World Bank database, although some data were taken from the IMF, and from the OECD database. We have originally provided a static snapshot of countries in our dataset, fixed to the third quarter of 2007, according to the way their macroeconomic data cluster in the map. In this way, we were able to capture the reciprocities and similarities from various emerging markets. In particular, looking at the results obtained it seems that the local proximity is the major contagion factor of economic vulnerabilities: neighbours are those with highest probability to get imbalanced when a critical event arises in a surrounding country. As second step, we have dynamically monitored countries' evolution path in the map over the time, with each variable observed from the first quarter of 1995 to the third quarter of 2007. Looking at the position changes of each country on the map along time, it has been possibile to infer the probability of future crises two years ahead.

It should be emphasized that the EWS model developed in this paper does certainly not constitute the final step towards a comprehensive EWS model. However, we believe that SOMs represent a new approach to EWS building, and hence the paper can address further steps towards developing more powerful tools for policy-makers in the future.

References

[1] Bussiere M, and M. Fratzscher (2002), Towards a new early warning system of financial crises. *European Central Bank working paper series*, 145, May 2002.

[2] Cattaneo Adorno, M. and M. Resta (2004), Reliability and convergence on Kohonen maps: an empirical study. In: M. Gh. Negoita, R. J. Howlett and L. C. Jain, *Proceedings of Knowledge-Based Intelligent Information and Engineering Systems*, Part I, Lecture Notes in Computer Science, Springer, 426-433.

[3] Cheng, Y. (1997), Convergence and ordering of Kohonen's batch map. *Neural Computation* 9, 1667-1676.

[4] De Bodt, E., Cottrell, M., and M. Verleysen (2002), Statistical tools to assess the reliability of self-organizing maps. *Neural Networks*, 15, 967-978.

[5] Jagtiani J., Kolari J., Lemieux C., and H. Shin (2003), Early Warning Models for Bank Supervision: Simpler Could Be Better. *Federal Reserve Bank of Chicago Economic Perspectives* (Quarter 3): 49Ű60.

[6] Kamin S., Schindler J., and S. Shawna (2001), The Contribution of Domestic and External Factors to Emerging Market Devaluation Crises: An Early Warning Systems Approach. *International Finance Working Paper* No. 711, Board of Governors of the Federal Reserve System.

[7] Kaminsky G. L., Lizondo S., and C. M. Reinhart (1998), Leading Indicators of Currency Crises. *IMF Staff Papers*, 45, No. 1, pp. 1-48.

[8] M. Y. Kiang, U. R. Kulkarni, M. Goul, A. Philippakis, R. T. Chi and E., Turban (1997), Improving the Effectiveness of Self-Organizing Map Networks Using a Circular Kohonen Layer. *Proc. Thirtieth Annual Hawaii Int. Conf. on System Sciences*.

[9] T. Kohonen (1982), *Self-Organizing Maps*, Springer Verlag, Berlin.

[10] C. Y. Liou and W. P. Tai (1999), Conformal self-organization for continuity on a feature map. *Neural Networks*, 12, 893-905.