Towards a Unified Statistical Framework to Evaluate Financial Crises Early Warning Systems

How to evaluate an EWS?

Bertrand Candelon[†], Elena-Ivona Dumitrescu[‡], Christophe Hurlin[‡]

[†]Maastricht University and [‡]University of Orléans

2009

Introduction

→ **From the subprime crisis to currency crises**

→ **Early Warning Systems (EWS) set up to ring before the occurence of crises**

Introduction

How can we specify an EWS model?

→ Rich literature (Kaminski et al. (1998), Kumar et al. (2003), Abiad (2003), etc.)

How can we evaluate the predictive abilities of an EWS?

→ Kaminski et al. (1998): signalling approach

- Threshold which minimizes the NSR criteria
- Type I and type II errors

→ Arbitrarely chosen cut-offs (Berg and Patillo (1999), Arias and Erlandsson (2005))

Originality

Our New EWS Evaluation Method

→ **I. Optimal cut-off**

→ **II. Credit-scoring evaluation criteria**
QPS, LPS, AUC, Pietra Index, Bayesian Error, Kuiper's score

→ **III. Comparison tests**

- Diebold-Mariano (1995) test for non-nested models
- Clark-West (2007) test for nested models
- Area under ROC comparison test

Contents

Step 1. A New EWS Evaluation Method

I. Optimal cut-off identification

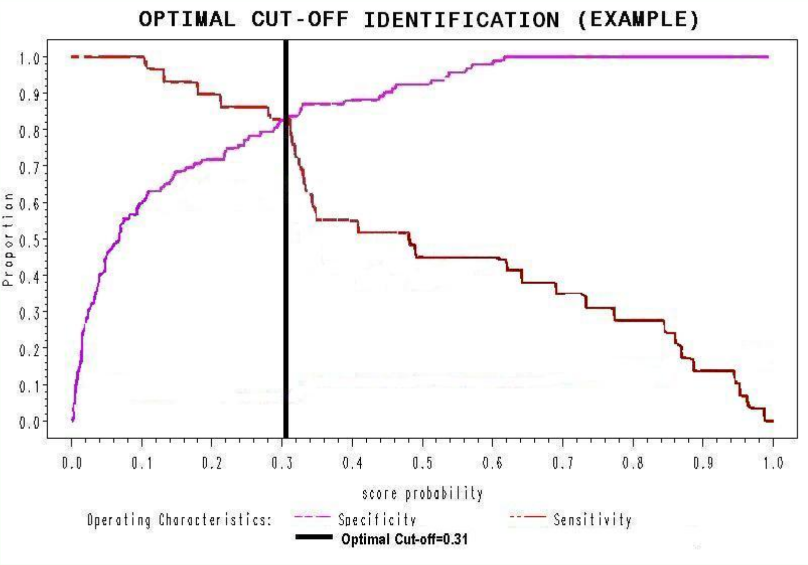
$$C^* = \text{Arg}_{\{C\}}[\textit{Sensitivity}(C) = \textit{Specificity}(C)], \text{ where } C \in [0, 1]$$

Definition 1.

Sensitivity is the number of crises correctly predicted for a cutoff C over the total number of crises in the sample

Definition 2.

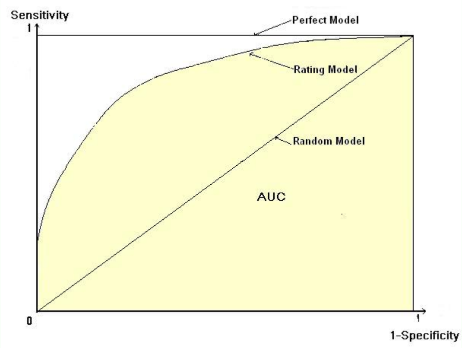
$1 -$ *Specificity* is the number of false alarms for a cutoff C over the total number of non-crises in the sample


OPTIMAL CUT-OFF IDENTIFICATION (EXAMPLE)
Proportion
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
score probability
Operating Characteristics:
Specificity
Sensitivity
Optimal Cut-off=0.31

II. Performance assessment criteria

The Area Under the ROC Curve and the Quadratic Probability Score

What is the ROC curve? (Receiving Operating Characteristic)



II. Performance assessment criteria

The Area Under the ROC Curve

$$A = \int_0^1 Sensitivity(1 - Specificity) d(1 - Specificity)$$

- Measure of the model's overall ability to discriminate between the cases correctly predicted and the false alarms
- For a perfect model AUC=1 while for a random one AUC=0.5

II. Performance assessment criteria

The Quadratic Probability Score

$$QPS = \frac{1}{T}\sum_{t=1}^{T} 2(\widehat{I_t} - I_t)^2$$

- Comparison of forecasts $(\widehat{I_t})$ and realizations (I_t)
- The closer QPS is to 0 the better the model is

III. Comparison tests

1. **Diebold-Mariano (1995) test for non-nested models**
2. **Clark-West (2007) test for nested models**
3. **Area under ROC comparison test (Delong et al. (1988))**

III. Comparison tests

Proposition 1: Let us denote by $M1$ and $M2$ two EWS models, and by $\widetilde{AUC_1}$ and $\widetilde{AUC_2}$ the associated areas under the ROC curve.

$$H0 : \widetilde{AUC_1} = \widetilde{AUC_2}$$

$$\frac{(\widetilde{AUC_1} - \widetilde{AUC_2})^2}{Var(\widetilde{AUC_1} - \widetilde{AUC_2})} \xrightarrow[T \to \infty]{d} \chi^2(1)$$

Step 2. EWS Specification and Estimation

To apply our evaluation methodology:

I. **Real crisis dating method (I_t)**

$\rightarrow$ KLR modified pressure index - Lestano and Jacobs (2004)
$\rightarrow$ The threshold equals two standard deviations above the mean

II. **Crisis probabilities ($\widehat{P}r_t$)**

$\rightarrow$ Panel logit with fixed effects
$\rightarrow$ Markov Switching Model with constant transition probabilities

I. Currency crisis dating method

KLR modified pressure index - Lestano and Jacobs (2004)

Definition 3. The 24 months crisis variable:

$$I_t = C24_{n,t} = \begin{cases} 1, & \text{if } \sum_{j=1}^{24} Crisis_{n,t+j} > 0 \\ 0, & \text{otherwise} \end{cases}$$

II. Empirical models

Model 1. Panel and time-series logit model

$$\Pr(C24_{nt} = 1) = \frac{\exp(\beta' x + f_n)}{1 + \exp(\beta' x + f_n)} \; \forall n \in \Omega_h,$$

where

- f_n represents the fixed effects
- x is the matrix of economic variables
- n is the country identifier
- Ω_h is the h^{th} cluster

Optimal country clusters: (Kapetanios procedure (2003))

II. Empirical models

Model 2. Markov model - Hamilton (1995)

$$KLRm_t = \mu_t(S_t) + \beta(S_t)x_t + \epsilon_t(S_t),$$

where

- $KLRm_t$ is the pressure index vector
- x_t represents the matrix of economic variables
- S_t follows a two states Markov chain

$$S_t = \begin{cases} 1, \text{if there is a crisis at time } t \\ 0, \text{if not} \end{cases}$$

II. Empirical models

Definition 4. The 24 months ahead forecasts (Arias and Erlandson (2005)):

$$\Pr(S_{t+1\ldots t+24} = 1|\Omega_t) = 1 - \Pr(S_{t+1\ldots t+24} = 0|\Omega_t)$$

$$= 1 - \{[P_{10}P_{00}^{(23)}\Pr(S_t = 1|\Omega_t)] + [P_{00}^{24}\Pr(S_t = 0|\Omega_t)]\},$$

- where P_{10} and P_{00} are elements of the transition probability matrix

II. Empirical models

From crisis probabilities to crisis forecasts

$$\widehat{I}_t = \begin{cases} 1, \textit{if } \Pr(C24_t = 1) > C^* \\ 0, \textit{otherwise} \end{cases} ,$$

where C^* is an **optimal cut-off** (see section 1)

Empirical Results

Empirical Results

I. Dataset
II. Optimal country clusters
III. Comparison tests
IV. Optimal model: cut-off identification and performance assessment criteria

I. Dataset

→ Monthly data in US dollars for the period 1985-2005 (6 Latin-American and 6 South-Asian Countries)

→ Market expectation (m.e.) variables:

- Yield spread
- Growth of stock market price index

→ Macroeconomic variables: Jacobs et al. (2003)

II. Optimal country clusters

Kapetanios procedure (2003)

1. Argentina, Brazil, Mexico, Venezuela
2. Peru, Uruguay
3. Korea, Malaysia, Taiwan
4. Philippines, Thailand
5. Indonesia

III. Comparison tests

Testing strategy

1. Logit with market-expectation variables vs. simple logit
2. Markov with market expectation variables and spread switching vs. Markov with market expectation variables
3. Best logit vs. best Markov specification

III.1. Logit with m.e. variables vs. simple logit

	ROC		Clark-West	
Country	test statistic	p-value	test statistic	pvalue
Argentina	0.0301	0.8622	0.1372	0.4454
Brazil	5.7105	**0.0169**	3.4901	**0.0002**
Indonesia	7.9917	**0.0047**	4.4332	**0.0000**
Korea	4.5357	**0.0332**	3.7746	**0.0001**
Malaysia	0.3859	0.5345	0.3288	0.3711
Mexico	$<$0.001	1.0000	0.6869	0.2460
Peru	0.0028	0.9577	2.1634	**0.0153**
Philippines	0.8738	0.3499	0.8709	0.1919
Taiwan	10.475	**0.0012**	3.5603	**0.0002**
Thailand	6.9801	**0.0082**	4.5964	**0.0000**
Uruguay	0.7443	0.3883	0.6656	0.2528
Venezuela	6.6647	**0.0098**	-2.0740	0.9810

$\ast$ The coefficients significant at a 5% level are in bold

III.2. Markov with m.e. variables and spread switching vs. Markov with m.e. variables

	ROC		Clark-West	
Country	test statistic	p-value	test statistic	pvalue
Argentina	10.930	**0.0009**	-6.7740	1.0000
Brazil	19.200	<**0.001**	8.0833	<**0.001**
Indonesia	36.319	<**0.001**	19.003	<**0.001**
Korea	4.8024	**0.0284**	-0.7131	0.7621
Malaysia	0.0064	0.9361	4.8475	<**0.001**
Mexico	0.0001	0.9930	-26.953	1.0000
Peru	6.9116	**0.0086**	9.7281	<**0.001**
Philippines	0.0906	0.7634	11.102	<**0.001**
Taiwan	0.5000	0.4795	1.4058	0.0799
Thailand	6.5530	**0.0105**	-7.7623	1.0000
Uruguay	111.15	<**0.001**	8.1857	<**0.001**
Venezuela	0.0691	0.7927	17.209	<**0.001**

$\ast$ The coefficients significant at a 5% level are in bold

III.3. Logit with m.e. variables vs. Markov with m.e. variables and spread switching

	ROC		Diebold-Mariano	
Country	test statistic	p-value	test statistic	pvalue
Argentina	62.678	<**0.001**	12.965	<**0.001**
Brazil	9.7859	**0.0018**	8.783	<**0.001**
Indonesia	46.529	<**0.001**	29.244	<**0.001**
Korea	9.8754	**0.0017**	12.207	<**0.001**
Malaysia	21.455	<**0.001**	17.066	<**0.001**
Mexico	17.829	<**0.001**	50.850	<**0.001**
Peru	45.942	<**0.001**	12.164	<**0.001**
Philippines	7.4266	**0.0064**	9.7129	<**0.001**
Taiwan	34.195	<**0.001**	16.591	<**0.001**
Thailand	45.902	<**0.001**	18.281	<**0.001**
Uruguay	125.00	<**0.001**	12.877	<**0.001**
Venezuela	17.351	<**0.001**	9.4665	<**0.001**

$\ast$ The coefficients significant at a 5% level are in bold

Comparison tests

Remarks

→ The panel logit model with market expectation variables works better than the Markov specifications

→ The introduction of market expectation variables has a positive effect on the forecasting performance of an EWS.

Best model - Optimal cut-off

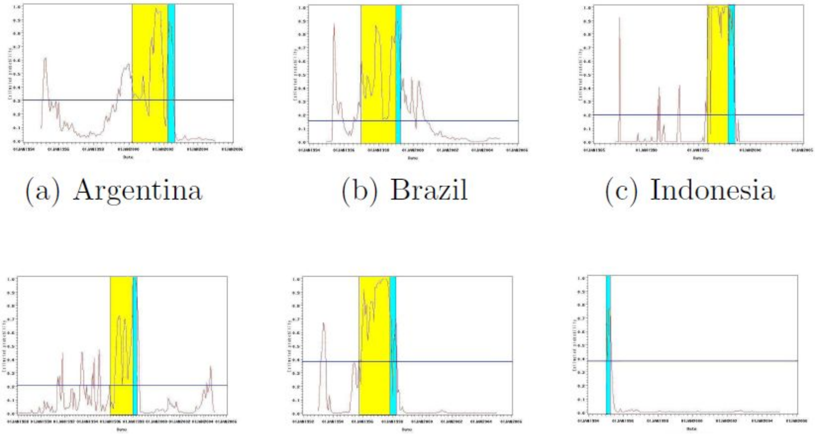
	Accuracy measures			Kaminski et al. (1998) NSR criteria		
Country	Cut-off	Sensitivity	Specificity	Cut-off	Sensitivity	Specificity
Argentina	**0.300**	82.76	82.61	0.620	41.38	100.0
Brazil	**0.160**	100.0	69.47	0.880	7.69	100.0
Indonesia	**0.200**	96.97	96.20	0.930	72.73	100.0
Korea	**0.206**	85.71	90.96	0.930	14.29	100.0
Malaysia	**0.380**	93.10	93.97	0.730	65.52	100.0
Mexico	**0.379**	100.0	99.15	0.390	75.00	100.0
Peru	**0.260**	100.0	82.72	0.940	12.90	100.0
Philippines	**0.346**	67.95	68.35	0.730	20.51	100.0
Taiwan	**0.160**	94.12	65.17	0.670	17.65	98.31
Thailand	**0.120**	90.32	61.29	0.321	25.81	96.24
Uruguay	**0.119**	93.33	75.73	0.900	50.00	100.0
Venezuela	**0.225**	85.71	67.90	0.330	64.29	77.78

- Optimal cut-off: $C \leq 0.38$
- Crisis and calm periods: correctly identified

Best model - Evaluation criteria

Country	AUC	Kuiper score	Pietra index	Bayesian error rate	QPS	LPS
Argentina	0.898	65.37	0.235	0.132	0.215	-0.325
Brazil	0.907	69.47	0.249	0.132	0.202	-0.311
Indonesia	0.996	93.17	0.330	0.0138	0.034	-0.058
Korea	0.920	76.67	0.273	0.0780	0.135	-0.228
Malaysia	0.985	87.07	0.311	0.048	0.083	-0.131
Mexico	0.998	99.15	0.350	0.008	0.011	-0.023
Peru	0.947	82.72	0.292	0.107	0.166	-0.266
Philippines	0.739	36.30	0.163	0.235	0.368	-0.558
Taiwan	0.739	36.30	0.163	0.235	0.368	-0.558
Thailand	0.811	51.61	0.192	0.138	0.218	-0.348
Uruguay	0.939	69.06	0.257	0.105	0.165	-0.246
Venezuela	0.777	53.61	0.189	0.257	0.370	-0.530

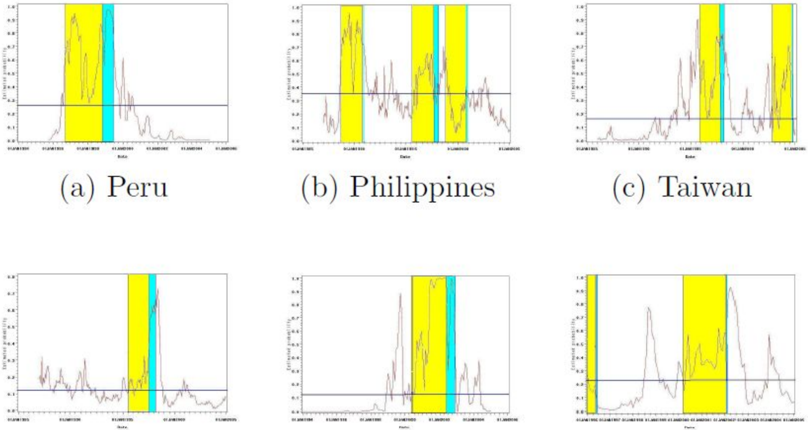
- Performance assessment criteria: close to the optimal values
- Robustness of the model to sensitivity analysis



(a) Argentina (b) Brazil (c) Indonesia

(d) Korea (e) Malaysia (f) Mexico

FIG. 1 – Predicted probability of crisis



(a) Peru (b) Philippines (c) Taiwan

(d) Thailand (e) Uruguay (f) Venezuela

FIG. 2 – Predicted probability of crisis (continued)

Conclusion

Conclusions

Objective: Developing a new EWS evaluation framework based on optimal cut-offs, credit-scoring criteria and comparison tests

→ Substantial improvement of the predictive power of EWS

→Markov models are not as efficient as panel logit model with market expectation variables

Conclusions

The optimal model

→ Predicts well most currency crises in the specified emerging markets

→ Robust to some sensitivity analysis

Extensions

→ Markov switching model with time varying probabilities

→ Other market expectation variables

→ A more consistent database (a longer period, more countries)

→ Out of sample validation