How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods

Published: 13 March 2012
Volume 60, pages 75–113, (2012) Cite this article

Bertrand Candelon, Elena-Ivona Dumitrescu & Christophe Hurlin

Abstract

This paper proposes an original and unified toolbox to evaluate financial crisis early-warning systems (EWS). It presents four main advantages. First, it is a model free method which can be used to assess the forecasts issued from different EWS (probit, logit, Markov switching models, or combinations of models). Second, this toolbox can be applied to any type of crisis EWS (currency, banking, sovereign debt, and so on). Third, it does not only provide various criteria to evaluate the (absolute) validity of EWS forecasts but also proposes some tests to compare the relative performance of alternative EWS. Fourth, this toolbox can be used to evaluate both in-sample and out-of-sample forecasts. Applied to a logit model for 12 emerging countries we show that the yield spread is a key variable for predicting currency crises exclusively for South-Asian countries. Besides, the optimal cut-off correctly allows us to identify now on average more than 2/3 of the crisis and calm periods.

Similar content being viewed by others

Notes

1. We do not tackle here the pertinence of the crisis dating. We assume that economic experts are able (ex-post) to precisely date the crisis periods. Nevertheless, a robustness analysis with respect to the potential inaccuracy of the crisis dating will be performed in the last section.

2. It can also be assumed that y_t equals one if a crisis occurs in a certain time horizon (6, 12, 24 months, and so on), so as to forecast the approximate timing of a crisis in some periods before it actually occurs (see KLR, Berg and Pattillo, 1999). This approach presents the advantage of giving the authorities the time necessary to implement appropriate policies to avoid an economic crash.

3. The *Log Probability Score* (*LPS*) corresponds to a loss function that penalizes large errors more heavily than *QPS*, with $LPS=-1/T\sum_{t=1}^{T}[(1-y_t)\ln(1-\hat{p}_t)+y_t\ln(\hat{p}_t)]$. This score ranges from 0 to ∞, with $LPS=0$ being perfect accuracy.

4. Theoretically, an alternative approach that jointly validates the optimal cut-off and the crisis probabilities may exist. Nevertheless, this approach is not feasible in our context. Indeed, the accuracy and misclassification measures cannot be employed, as they have been used to identify the optimal cut-off and no other adequate measures have been proposed so far.

5. Also called Correct Classification Frontier, as in Jorda, Moritz, and Taylor (2011).

6. This nonparametric estimator of the AUC criterion has recently been considered by Jorda, Moritz, and Taylor (2011) in the EWS literature, so as to compare different specifications with the random model (AUC=0.5).

7. Contrary to Jorda, Moritz, and Taylor (2011), who rely on a graphical comparison of the AUC for different models, we develop a statistical framework to evaluate EWS.

8. Let us assume that model 1 is the parsimonious model and model 2 is the larger one that reduces to model 1 if some of its parameters are set to 0. The corrected statistic proposed by Clark and West (2007), denoted *CW*, is defined as follows:

$$CW=\frac{\sqrt{T}\bar{f}}{\sigma_{\bar{f},0}}\xrightarrow[T\to\infty]{d}N(0,1), \tag{16}$$

where $\hat{f}_t=(y_t-\hat{p}_{1,t})^2-[(y_t-\hat{p}_{1,t})^2-[(y_t-\hat{p}_{2,t})^2-(\hat{p}_{2,t}-\hat{p}_{1,t})^2]$, $\bar{f}$ is the sample average of $\hat{f}_t$ and $\sigma_{\bar{f},0}^{2}$ is the sample variance of $\hat{f}_t-\bar{f}$.

9. Note that, as a robustness check, we have also considered the pooled logit model as well as the optimal clusters derived from the Kapetanios procedure.

10. Argentina, Brazil, Mexico, Peru, Uruguay, Venezuela, Indonesia, South Korea, Malaysia, Philippines, Taiwan, and Thailand.

11. Berg and Cooke (2004) show that considering a forecast horizon larger than 1 leads to autocorrelation in the crisis variable. This stylized fact is confirmed by Harding and Pagan (2006).

12. The results for the other countries are available on request.

13. In the case of KLR the threshold equals three standard deviations; however, in this case, Taiwan would never register any currency crises, which is historically not accurate. For example, Taiwan was not exempted from the Asian crisis in 1997.

References

Abiad, Abdul, 2003, “Early Warning Systems: A Survey and a Regime Switching Approach,” IMF Working Paper.

Arias, Guillaume, and Ulf G. Erlandsson, 2005, “Improving Early Warning Systems with Markov Switching Model,” C.E.F.I. Working Paper, 0502.

Bai, Jushan, and Serena, Ng, 2001, “A New Look at Panel Testing of Stationarity and the PPP Hypothesis,” Boston College Working Papers in Economics 518 (Boston College Department of Economics).

Barrios, Salvador, Per Iversen, Magdalena Lewandowska, and Ralph Setzer, 2009, "Determinants of Intra-Euro Area Government Bond Spreads During the Financial Crisis," Economic Papers 388.

Basel Committee on Banking Supervision, 2005, "Studies on the Validation of Internal Rating Systems," Working Paper no. 14 (Bank for International Settlements).

Berg, Andrew, and Catherine A. Pattillo, 1999, "Predicting Currency Crises: The Indicators Approach and an Alternative," Journal of International Money and Finance, Vol. 18, pp. 561–586.

Article Google Scholar

Berg, Andrew, and Rebecca N. Cooke, 2004, "Autocorrelation Corrected Standard Errors in Panel Probits: An Application to Currency Crisis Prediction," IMF Working Paper.

Berg, Jeroen van den, Bertrand Candelon, and J.P. Jean-Pierre Urbain, 2008, "A Cautious Note on the Use of Panel Models to Predict Financial Crises," Economics Letters, Vol. 101, No. 1, pp. 80–83.

Article Google Scholar

Bordo, Michael D., Barry Eichengreen, Daniela Klingebiel, and Maria-Soledad Martinez-Peria, 2001, "Is the Crisis Problem Growing More Severe?" Economic Policy, Vol. 32, pp. 51–82.

Google Scholar

Bussiere, Matthieu, and Michael Fratzscher, 2006, "Towards a New Early Warning System of Financial Crises," Journal of International Money and Finance, Vol. 25, No. 6, pp. 953–973.

Article Google Scholar

Clark, Todd E., and Kenneth D. West, 2007, “Approximately Normal Tests for Equal Predictive Accuracy in Nested Models,” Journal of Econometrics, Vol. 138, No. 1, pp. 291–311.

Article Google Scholar

Clark, Todd E., and Michael W. McCracken, 2001, “Tests of Equal Forecast Accuracy and Encompassing for Nested Models,” Journal of Econometrics, Vol. 105, No. 1, pp. 85–110.

Article Google Scholar

DeLong, Elizabeth R., Dwight M. DeLong, and Daniel L. Clarke-Pearson, 1988, “Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach,” Biometrics, Vol. 44, No. 3, pp. 837–845.

Article Google Scholar

Diebold, Francis X., and Glenn D. Rudebusch, 1989, “Scoring the Leading Indicators,” The Journal of Business, Vol. 62, No. 3, pp. 369–391.

Article Google Scholar

Diebold, Francis X., and Roberto S. Mariano, 1995, “Comparing Predictive Accuracy,” Journal of Business and Economic Statistics, Vol. 13, No. 3, pp. 253–263.

Google Scholar

Engelmann, Bernd, Evelyn Hayden, and Dirk Tasche, 2003, “Testing Rating Accuracy?” Risk, Vol. 16, pp. 82–86.

Google Scholar

Estrella, Arturo, and Frederic S. Mishkin, 1996, "The Yield Curve as a Predictor of US Recessions," Current Issues in Economics and Finance, Vol. 2, No. 7, pp. 1–5.

Google Scholar

Estrella, Arturo, and Gikas A. Hardouvelis, 1991, "The Term Structure as a Predictor of Real Economic Activity," The Journal of Finance, Vol. 46, No. 2, pp. 555–576.

Article Google Scholar

Estrella, Arturo, and Mary R. Trubin, 2006, "The Yield Curve as a Leading Indicator: Some Practical Issues," Current Issues in Economics and Finance, Vol. 12, No. 5, pp. 1–7.

Google Scholar

Frankel, Jeffrey, and George Saravelos, (2010), "Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008–09 Global Crisis," NBER Working Paper 16047.

Fratzscher, Marcel, 2003, "On Currency Crises and Contagion," International Journal of Finance and Economics, Vol. 8, No. 2, pp. 109–129.

Article Google Scholar

Fuertes, Ana-Maria, and Elena Kalotychou, 2007, "Optimal Design of Early Warning Systems for Sovereign Debt Crises," International Journal of Forecasting, Vol. 23, No. 1, pp. 85–100.

Article Google Scholar

Guidolin, Massima, and Yu M. Tam, 2010, "A Yield Spread Perspective on the Great Financial Crisis: Break-Point Test Evidence," Federal Reserve Bank of St.

Louis Working Paper No. 26.

Gould, William ., Pitblado Jeffrey, and Sribney, William, and Stata Corporation, 2005, “Maximum Likelihood Estimation with Stata,” (College Station, TX: Stata Press).

Harding, Don, and Adrian R. Pagan, 2006, “The Econometric Analysis of Constructed Binary Time Series,” Working Papers Series 963 (The University of Melbourne).

Hoeffding, Wassily, 1948, “A Class of Statistics with Asymptotically Normal Distributions,” Annals of Statistics, Vol. 19, pp. 293–325.

Article Google Scholar

Im, Kyung S., Hashem M. Pesaran, and Shin Yongcheol, 1997, “Testing for Unit Roots in Heterogenous Panels,” DAE, Working Paper 9526 (University of Cambridge).

Jacobs, Jan P.A.M., Gerard H. Kuper, and Lestano, 2004, “Financial Crisis Identification: A Survey,” Working Paper (University of Groningen).

Jacobs, Jan P.A.M., Gerard H. Kuper, and Lestano, 2008, “Currency Crises in Asia: A Multivariate Logit Approach,” in International Finance Review Asia Pacific Financial Markets: Integration, Innovation and Challenges, Vol. 8, eds. by S.J. Kim and M. McKenzie, pp. 157–173.

Jorda, Oscar, Schularick Moritz, and Alan M. Taylor, 2011, “Financial Crises, Credit Booms, and External Imbalances: 140 Years of Lessons,” IMF Economic Review, Vol. 59, No. 2, pp. 340–378.

Article Google Scholar

Kaminsky, Gracilea L., 2003, “Varieties of Currency Crises,” NBER Working Paper No. 10193.

Kaminsky, Graciela L., Saul Lizondo, and Carmen Reinhart, 1998, “Leading Indicators of Currency Crises,” IMF Staff Papers, Vol. 45, No. 1, pp. 1–48.

Article Google Scholar

Kapetanios, George, 2003, “Determining the Poolability of Individual Series in Panel Datasets,” Working Paper 499.

Kraft, Holger, Gerald Kroisandt, and Marlene Muller, 2004, “Redesigning Ratings: Assessing the Discriminatory Power of Credit Scores Under Censoring,” Working Paper.

Kumar, Mohan, Uma Moorthy, and William Perraudin, 2003, “Predicting Emerging Market Currency Crashes,” Journal of Empirical Finance, Vol. 10, pp. 427–454.

Article Google Scholar

Kydland, Finn E., and Edward C. Prescott, 1991, “The Econometrics of the General Equilibrium Approach to Business Cycles,” Scandinavian Journal of Economics, Vol. 93, No. 2, pp. 161–178.

Article Google Scholar

Lambert, Jerome, and Ilya Lipkovich, 2008, “A Macro for Getting More Out of Your ROC Curve,” SAS Global forum, paper 231.

Lau, Francis, Sunny Yung, and Ivy Yong, 2003, “Introducing a Framework to Measure Resilience of an Economy,” Hong Kong Monetary Authority Quarterly

Bulletin.

Lestano, Jacobs, P.A.M. Jan, and Gerard H. Kuper, 2003, “Indicators of Financial Crises Do Work! An Early-Warning System for Six Asian Countries,” Working Paper.

Martinez-Peria, Maria-Soledad, 2002, “A Regime-Switching Approach to the Study of Speculative Attacks: A Focus on EMS Crises,” Empirical Economics, Vol. 27, No. 2, pp. 299–334.

Article Google Scholar

Maddala, Gangadharrao S., and Shaowen Wu, 1999, “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test,” Oxford Bulletin of Economics and Statistics, Special Issue, pp. 631–652.

Mitchener, Kris James, and Marc D. Weidenmier, 2006, “The Baring Crisis and the Great Latin American Meltdown of the 1890s,” NBER Working Paper No. 13403.

Pesaran, Hashem M., 2003, “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence,” Mimeo (University of Southern California).

Renault, Olivier, and Arnaud De Servigny, 2004, The Standard & Poor's Guide to Measuring and Managing Credit Risk (New York: McGraw-Hill, 1st ed.).

Google Scholar

Rose, Andrew K., and Mark M. Spiegel, 2010, “Cross-Country Causes and Consequences of The 2008 Crisis: International Linkages and American Exposure,” Pacific Economic Review, Vol. 15, No. 3, pp. 340–363.

Article Google Scholar

Rose, Andrew K., and Mark M. Spiegel, 2011, “Cross-Country Causes and Consequences of the Crisis: An Update,” European Economic Review, Vol. 55, No. 3, pp. 309–324.

Article Google Scholar

Stein, Roger M., 2005, “The Relationship Between Default Prediction and Lending Profits: Integrating ROC Analysis and Loan Pricing,” Journal of Banking and Finance, Vol. 29, pp. 1213–1236.

Article Google Scholar

Williams, Rick L., 2004, “A Note on Robust Variance Estimation for Cluster-Correlated Data,” Biometrics, Vol. 56, No. 2, pp. 645–646.

Article Google Scholar

Wright, Jonathan H., 2006, “The Yield Curve and Predicting Recessions,” Finance and Economic Discussion Series No. 7 (Federal Reserve Board).

Zhang, Zhiwei, 2001, “Speculative Attacks in the Asian Crisis,” IMF Working Paper 189 (Washington, DC: International Monetary Fund).

Additional information

Supplementary Information accompanies the paper on *IMF Economic Review* website (http://www.palgrave.com/imfer)

*Bertrand Candelon is Professor of International Monetary Economics at Maastricht University (the Netherlands), Christophe Hurlin is Professor of Economics at the University of Orleans (France) and Elena-Ivona Dumitrescu is

PhD student at both Maastricht and Orleans universities. The authors thank two anonymous referees as well as Pierre-Olivier Gourinchas, the editor of the *IMF Economic Review* for stimulating comments. The authors are also indebted to the participants at the IMF institute Seminar, the 2010 European Economic Association Congress (Glasgow), the 2010 Econometric Society World Meeting (Shanghai), the 2010 meeting of the Association Francaise de Sciences Economiques (Paris), the 2010 business cycle meeting at Eurostat (Luxemburg) and the 2009 MIFN meeting (Luxemburg) for their questions and reactions. The usual disclaimers apply.

Electronic supplementary material

Supplementary Tables (download DOC)

Appendices

Appendix I

A.I. Comparison of *ROC* Curves Test

The nonparametric *test of comparison of ROC curves* has been proposed by DeLong, DeLong, and Clarke-Pearson (1988). It is based on the comparison of the areas under the ROC curves associated with the two EWS models, denoted AUC_1 and AUC_2. The null of the tests corresponds to the equality of areas under the *ROC* curves, that is H_0: $AUC_1 = AUC_2$. The test statistic is defined as:

$$W_{AUC} = \frac{(AUC_1 - AUC_2)^2}{\mathbb{V}(AUC_1 - AUC_2)}. \qquad \text{(A.1)}$$

Under the null, it has an asymptotic $\chi^2(1)$ distribution. By definition the asymptotic variance of the difference ($AUC_1 - AUC_2$) is equal to:

$$\mathbb{V}(AUC_1 - AUC_2) = \mathbb{V}(AUC_1) + \mathbb{V}(AUC_2) - 2cov(AUC_1, AUC_2). \qquad \text{(A.2)}$$

Each of these three elements can be estimated using a nonparametric kernel estimator. Let us consider $\mathbb{V}$ the variance-covariance matrix of the vector (AUC_1

$AUC_2)'$. A nonparametric kernel estimator of $\mathbb{V}$, denoted $\hat{\mathbb{V}}$, can be derived from the theory developed for generalized U-statistics by Hoeffding (1948) and Mann-Whitney statistics. Formally, we have:

$$\underset{(2,2)}{\hat{\mathbb{V}}} = (T_1)^{-1} \underset{(2,2)}{\hat{S}_1} + (T_0)^{-1} \underset{(2,2)}{\hat{S}_0}, \tag{A.3}$$

where T_1 (respectively T_0) is the number of crisis (respectively calm) periods in the sample, and $\hat{S}_1$ (respectively $\hat{S}_0$) denotes the estimated variance for the crisis (respectively calm) periods.

$$\hat{S}_1 = \frac{1}{T_0^2(T_1 - 1)} \sum_{i:y_i=1} \begin{pmatrix} \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_1\right]^2 & \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_1\right] \times \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_2\right] \\ \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_1\right] \times \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_2\right] & \left[\sum_{j:y_j=0} K(\hat{p}_j, \hat{p}_i) - T_0 \times AUC_2\right]^2 \end{pmatrix}. \tag{A.4}$$

Similarly, we have:

$$\hat{S}_0 = \frac{1}{T_1^2(T_0-1)} \sum_{j:y_j=0} \begin{pmatrix} \left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_1\right]^2 & \left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_1\right] \\ \left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_1\right] & \times\left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_2\right] \\ \times\left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_2\right] & \left[\sum_{i:y_i=1} K(\hat{p}_j,\hat{p}_i) - T_1 \times AUC_2\right]^2 \end{pmatrix}, \tag{A.5}$$

where $K(.)$ denotes a kernel function of the estimated crisis probabilities in crisis periods ($y_i = 1$) and calm periods ($y_j = 0$) defined by:

$$K(\hat{p}_j,\hat{p}_i) = \begin{cases} 1 & \text{if } \hat{p}_i < \hat{p}_j \\ \frac{1}{2} & \text{if } \hat{p}_i = \hat{p}_j \\ 0 & \text{if } \hat{p}_i > \hat{p}_j. \end{cases} \tag{A.6}$$

Appendix II

B.I. Data Set

There is no official currency crisis dating method similar to the one NBER proposes for recessions. Therefore, a crisis episode is generally detected when an index of speculative pressure exceeds a certain threshold. Many alternative indices have been developed and used for identifying currency crises. But they are all nonparametric termination rules that take into consideration the size of the movements in a combination of a number of series. Lestano and Jacobs (2004) compare several currency crisis dating methods, aiming to identify the one that recognizes most of the crises categorized by the IMF for the 1997 Asian flu. They conclude that the KLR modified index, the Zhang original index (Zhang, 2001), and extreme values applied to the KLR modified index perform best.

Following their results, we identify crisis periods using the KLR modified pressure index (KLRm), which, unlike the KLR index, also includes interest rates:

$$KLRm_{it} = \frac{\Delta e_{it}}{e_{it}} - \frac{\sigma_e}{\sigma_r}\frac{\Delta r_{it}}{r_{it}} + \frac{\sigma_e}{\sigma_{ir}}\Delta ir_{it}, \tag{A.7}$$

where e_{it} denotes the exchange rate (that is, units of country i's currency per U.S. dollar in period t), $r_{n,t}$ represents the foreign reserves, while ir_{it} is the interest rate. Meanwhile, the standard deviations σ_X are actually the standard deviations of the relative changes in the variables, $\sigma_{(\Delta X_{it}/X_{it})}$, where X denotes each variable separately, including the exchange rate and the foreign reserves, with $\Delta X_{it} = X_{it} - X_{i,t-6}$. For the interest rate, σ_{ir} is the standard deviation of the absolute changes in interest rate. For both subsamples, the threshold equals two standard deviations above the mean:[13]

$$Crisis_{it} = \begin{cases} 1 & \text{if } KLRm_{it} > 2\sigma_{KLRm_{it}} + \mu_{KLRm_{it}} \\ 0 & \text{otherwise.} \end{cases} \tag{A.8}$$

To check the robustness of our results to the dating method, we also consider the *Zhang* pressure index instead of the *KLRm*. It is defined as follows:

$$Crisis_{it} = \begin{cases} 1 & \text{if } \left(\begin{array}{ll} \frac{\Delta e_{it}}{e_{it}} > \beta_1 \sigma'_{e_{it}} + \mu_{e_{it}} & \text{or} \\ \frac{\Delta r_{n,t}}{r_{it}} < \beta_2 \sigma'_{r_{it}} + \mu_{r_{it}} & \end{array} \right. \\ 0 & \qquad \text{otherwise.} \end{cases} \tag{A.9}$$

where σ_{eit}' is the standard deviation of $(\Delta e_{it}/e_{it})$ in the sample of $(t-36, t-1)$, and σ_{rit}' is the standard deviation of $(\Delta r_{it}/r_{it})$ in the sample of $(t-36, t-1)$. The thresholds are set to $\beta_1=3$ and $\beta_2=-3$. Contrary to the KLRm index, the interest rates are excluded from the ZCC and the thresholds used are time-varying for each component.

From a macroeconomic point of view, it is more important to know if there will be a crisis in a certain horizon than in a certain month, because this time period allows the state to take steps to prevent the crisis. Consequently, we define for each country $C24_t$, which corresponds to y_t from our general framework and thus serves as the crisis dummy variable taking the value of 1 if there will be a crisis in the following 24 months and 0 otherwise:

$$C24_{it} = \begin{cases} 1 & \text{if } \sum_{j=1}^{24} Crisis_{it+j} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{A.10}$$

At the same time, several explanatory variables from three economic sectors are considered (Lestano and Kuper, 2003) on a monthly frequency and denoted in U.S. dollars:

1 *External sector*: the one-year growth rate of international reserves, the one-year growth rate of imports, the one-year growth rate of exports, the ratio of M2 to foreign reserves, and the one-year growth rate of M2 to foreign reserves.

2 *Financial sector*: the one-year growth rate of M2 multiplier, the one-year growth rate of domestic credit over GDP, the one-year growth rate of real bank deposits, the real interest rate, the lending rate over deposit rate, and the real interest rate differential.

3 *Domestic real and public sector*: the industrial production index.

As in Kumar, Moorthy, and Perraudin (2003), we reduce the impact of extreme values by using the formula: $f(x_t) = sign(x_t) \times ln(1+|x_t|)$. Traditional first-generation (Im, Pesaran, and Shin, 1997 and Maddala and Wu, 1999) and second-generation (Bai and Ng, 2001 and Pesaran, 2003) panel unit root tests are performed, leading to the rejection of the null hypothesis of stochastic trend except for the lending rate over deposit rate and industrial production index indicators. Hence, these series are substituted by their first differences.

Finally, we identify the most correlated leading indicators for each country. Two indicators are considered as being correlated for a certain country if Pearson's correlation coefficient is higher than a 30 percent threshold. It seems that growth of real exchange rate and real interest rate are highly correlated with most indicators for all countries, whereas the first difference of lending rate over deposit rate, the first difference of the industrial production index, and yield spread are the least correlated ones with all the other indicators for all the 12 countries. The competing models are defined such that no couple of indicators is correlated in more than four countries. We hence identify the leading indicators by minimizing the AIC and BIC information criteria of the pooled panel data models, that is, growth of international reserves, growth of exports, growth of domestic credit over GDP, first difference of lending over deposit rate, first difference of industrial production index, and yield spread. The missing values through the series are replaced using cubic splines interpolation, but when the series revealed missing values at the beginning of the sample, such as “the one-year growth of terms of trade” or “yield spread,” the corresponding observations are dropped from the analysis, leading to an unbalanced panel framework. Table 8 shows the period covered by the leading indicators for each of the 12 countries.

Table 8 Database

Appendix III

C.I. A Robust Estimator of the Variance of the Parameters

To compute robust estimators of the variance for logit models we use a sandwich estimator. Technically, the variance-covariance matrix of the estimators is asymptotically equal to the inverse of the hessian matrix: $\mathbb{V}(\hat{\beta}) = -H(\hat{\beta})^{-1}$ However, this is appropriate only if we employ the real Data Generating Process (DGP). For a more permissive method from this point of view, we define the variance vector as follows:

$$\mathbb{V}(\hat{\beta}) = (-H(\hat{\beta})^{-1})\mathbb{V}(g(\hat{\beta}))(-H(\hat{\beta})^{-1}), \qquad \text{(A.11)}$$

where $H(\beta)^{-1}$ is the inverse of the hessian matrix, and $\mathbb{V}(g(\hat{\beta}))$ is the variance of the gradient. Using the empirical variance estimator of the gradient we find that:

$$\mathbb{V}(\hat{\beta}) = -T/(T-1)H(\hat{\beta})^{-1}\left\{\sum_{t=1}^{T} g_t(\hat{\beta})g_t(\hat{\beta})'\right\}(-H(\hat{\beta})^{-1}), \qquad \text{(A.12)}$$

which is a robust variance estimator for the time-series model.

The main advantage of this sandwich method is that it can also be applied in the case of grouped data, as in our case. It is important to note that in the current situation, each country from a cluster is a group of time-series observations that are correlated. Thus, the observations corresponding to a country are not treated as independent, but rather the countries themselves which form the clusters, are considered independent. Therefore, instead of using $g_t(\beta)$, we use the sum of $g_t(\beta)$ for each country, while T is replaced by the number of countries in a cluster. These changes ensure the independence of so-called superobservations entering the formula (Gould and others, 2005).

Rights and permissions

About this article

Cite this article

Candelon, B., Dumitrescu, EI. & Hurlin, C. How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods. *IMF Econ Rev* **60**, 75–113 (2012). https://doi.org/10.1057/imfer.2012.4

Published
13 March 2012

Issue date
01 April 2012

DOI
https://doi.org/10.1057/imfer.2012.4

JEL Classifications

C33 F37