Business School

WORKING PAPER SERIES

Working Paper

2014-161

Currency Crisis Early Warning Systems: Why They should be Dynamic

Bertrand Candelon
Christophe Hurlin
Elena Dumitnescu

<http://www.ipag.fr/fr/accueil/la-recherche/publications-WP.html>

IPAG Business School
184, Boulevard Saint-Germain
75006 Paris
France

IPAG working papers are circulated for discussion and comments only. They have not been peer-reviewed and may not be reproduced without permission of the authors.

Currency Crisis Early Warning Systems: Why They should be Dynamic

October 2013

Abstract

Traditionally, financial crisis Early Warning Systems (EWSs) rely on macroeconomic leading indicators to forecast the occurrence of such events. This paper extends such discrete-choice EWSs by taking into account the persistence of the crisis phenomenon. The dynamic logit EWS is estimated using an exact maximum likelihood estimation method both in a time series and panel form. This model's forecasting abilities are then scrutinized by using an evaluation methodology recently designed specifically for EWSs. When applied for predicting currency crises for 16 countries, this new EWS turns out to exhibit significantly better predictive abilities than the existing static one, both in- and out-of -sample, thus supporting the use of dynamic specifications for EWSs for financial crises.

Key words: dynamic models, currency crisis, Early Warning System.

J.E.L. Classification: C33, F37

The authors thank the editor, an associate editor, as well as two anonymous referees for useful comments on the earlier version of the paper as well as the participants at the 6th Eurostat Colloquium on "Modern Tools for Business Cycle Analysis" in Luxembourg at the 4th annual MIFN Workshop in Shandong University, at the internal seminars of the CORE-Université Catholique de Louvain, IMF research department and the Central Bank of Tunisia. The usual disclaimers apply.

1 Introduction

The recent subprime crisis has renewed the interest of both policymakers and academics in Early Warning Systems (EWSs). These models are constructed so that they should be able to set off the alarm before the occurrence of a financial crisis (banking, currency, debt, etc.), thereby providing authorities with enough time to implement adequate rescue policies to prevent or at least to attenuate the perverse effects of the turmoil. Unfortunately, existing EWSs remained silent at the brink of the recent financial crisis, leading researchers to call into question their usefulness (see Rose and Spiegel, 2010). One of the reasons behind this is that EWS models are fundamentally static in the sense that they explain crisis occurrence only through changes in exogenous, macroeconomic variables. This paper contributes to the existing literature on financial crises, and in particular the literature on currency crises, by developing a new generation of EWSs, which takes into account the endogenous dynamic of this phenomenon.

The objective of an EWS model consists in accurately detecting the occurrence of a crisis, which is represented by a binary variable which takes the value of one when the event occurs, and the value of zero otherwise. Given the dichotomic character of the crisis indicator, traditional times series econometrics techniques cannot be directly implemented. In this context, the first approach to construct an EWS was based upon a signaling approach (Kaminski, Lizondo and Reinhart, 1998). Using a large set of potentially informative variables, Kaminsky et al. (1998) identified thresholds beyond which a crisis is signaled by each of these leading indicators. The properties of such an EWS clearly depend on these cut-off points. Kaminsky et al. (1998) estimated each of these thresholds as the value that minimizes the ratio between the number of crises missed and correctly detected, also called the noise-to-signal ratio.¹ Once the variable specific thresholds are determined, an aggregate indicator of an impending crisis is constructed as a weighted combination of these variables. The weights correspond to the inverse of the associated noise to signal ratio. As the economy approaches a crisis period, this EWS should exhibit a positive trend, *i.e.* indicate an increase in the probability of occurrence of a crisis.

A second class of EWS models began to emerge with work by Berg and Patillo (1999). These authors argue that it is impossible to use statistical methods to make inferences in the signaling framework and opt for a static panel probit EWS. In their framework, the binary crisis indicator is treated as an endogenous variable which is explained by a set of macroeconomic variables. Furthermore, it becomes possible to formally test the significance of leading indicators of currency crises. Evaluation criteria such as the quadratic probability score (QPS) and the log probability score (LPS) indicate that their EWS exhibits better forecasting abilities (in- and out-of-sample) than the EWS proposed by Kaminski et al. (1998).

Given that the estimation method for panel discrete-choice variables is quite standard and available in almost all econometric software, the panel EWS framework has been extensively implemented in applied studies. Several extensions have been proposed. Kumar et al. (2003) advocate the use of panel logit models instead of panel probit. Bussiere and Fratzscher (2006) suppose that a post-crisis specific regime may occur and define the crisis indicator as a variable with three states instead of two, thus estimating a multinomial logit EWS. Falcetti and Tudela (2006) estimate a binary model in a maximum-simulated likelihood framework including the lagged crisis indicator on the right-hand side. Berg et al. (2008) analyse the presence of country clusters, *i.e.* groups, and their consequences for an EWS.

One feature that the above-mentioned EWSs have in common is that they are static², *i.e.* they assume

¹Alternative estimation methods of this threshold are also available (see Candelon et al., 2012).

²An alternative strategy to build EWSs by integrating crisis dynamics has been proposed via Markov-switching models (see Abiad, 2003 *inter alii*). But these models rely on continuous market pressure indices (weighted average of macroeconomic variables that indicate balance-of-payment vulnerability) and do not

that the probability of crisis depends only on a set of macroeconomic variables which reflect the economic policies implemented. This assumption is not supported by empirical studies, which show that the longer a country is in a crisis period the higher the probability that it will exit the crisis, whatever the political reaction is (see Tudela, 2004). Neglecting the impact of endogenous crisis persistence may lead to an incorrect evaluation of the weight of leading macroeconomic indicators, and hence to the implementation of inadequate policies.

The present paper proposes a new generation of EWSs which reconcile the binary-choice property of the crisis variable and the persistence dimension of the crisis process. The specifications considered include both macroeconomic variables, *i.e.* the source of exogenous crisis persistence, and the dynamics of the crisis, *i.e.* the source of endogenous persistence.

The contribution of this paper to the EWS literature is twofold: first we propose new dynamic EWS models, and second, we accurately scrutinize their forecasting abilities. Intuitively, accounting for crisis dynamics may improve the fit of the model, as crises are persistent events (they last more than one period). But do these new EWSs exhibit better forecasting abilities than existing models? To answer this question, we gauge both the in-sample and out-of-sample forecasting performance of static and dynamic discrete-choice EWSs. As argued in Candelon et al. (2012), most of the EWS literature focuses on identifying which leading indicators should be included in such a model while the evaluation of EWSs is generally based on ‘second-best’ criteria, *e.g.* noise-to-signal ratio, QPS, LPS. We are the first to rely on the recent evaluation methodology proposed by Candelon et al. (2012), *i.e.* the area under the *ROC* curve (*AUC*) criterion and the associated *ROC* test, in our quest for the best-performing discrete-choice EWS.

The dynamic EWS framework we propose has several advantages. First, it allows us to easily estimate and compare several binary EWS specifications including the static one. To do this, we implement the exact maximum likelihood estimation method proposed by Kauppi and Saikkonen (2008).³ Second, it can be applied to construct EWSs for any type of crisis, not only currency crises as in our application. Furthermore, we extend this time-series framework to a dynamic fixed effects panel EWS including the score correction proposed by Carro (2007).⁴

As will emerge in what follows, the dynamic model including the lagged binary dependent variable which we propose outperforms (both in-sample and out-of sample) the specifications considered in the previous EWSs, in particular the traditional static logit. Most importantly, our new EWS has good out-of-sample forecasting abilities not only one-period-ahead, but also for a longer 6-month forecast horizon. This is an important advantage from the policymaker’s point of view, as an EWS is supposed to give regulators enough time to take steps to prevent a crisis or at least reduce its intensity. The dynamic EWS we propose correctly identifies most of the out-of-sample crisis and calm periods for most of the countries examined, thus supporting the use of dynamic specifications in the quest for an optimal currency-crisis EWS model.

This paper is structured as follows. The new estimation methods used for dynamic binary-choice EWSs

consider the binary crisis indicator as the dependent variable. Since the superior forecasting abilities of static discrete-choice EWSs relatively to Markov-switching ones have been emphasized in Candelon et al. (2009), we do not consider Markov-switching EWSs in this paper.

³Kauppi and Saikkonen (2008) implement this estimation method for business cycle analysis and, to our knowledge, it has not been considered in the financial crises literature to date. The only existing dynamic EWS based on a discrete-choice specification has been proposed by Falcetti and Tudela (2006). The authors consider the case where the previous regime can affect the current crisis probability and use a less straightforward maximum simulated likelihood estimation method. In contrast, our framework allows for different specifications of the dynamics of the crisis. It can be easily written in most common econometric software and it is not time-consuming (results are obtained in a few seconds).

⁴All Matlab codes are available from the authors upon request.

both in time-series and panel are discussed in section 2. The database, the currency crisis dating method and a preliminary in-sample analysis are presented in section 3. Section 4 proposes both an in-sample and an out-of-sample comparison of the forecasting abilities of the models, while section 5 concludes.

2 A Dynamic Specification of EWSs

To date, almost all EWS models are static and do not exploit the persistence property of crises captured by lagged endogenous indicators. From an econometric perspective, this issue is known as the debate on the exogenous/endogenous persistence of a phenomenon. This paper is hence the first to consider a dynamic version of EWSs based on an exact maximum likelihood estimation method by relying on Kauppi and Saikkonen (2008).

2.1 Specification and Estimation

Let us consider first the time-series version of the dynamic binary-choice EWS. We denote by $\{y_{i,t}\}_{t=1}^T$ the currency crisis binary variable for country i , $i \in \{1, 2, \dots, N\}$, which takes the value 1 during crisis periods and 0 otherwise. Similarly, $\{x_{i,t}\}_{t=1}^T$ represents the matrix of explanatory variables. For ease of notation, hereafter the country index i will be omitted.

The one-step-ahead dynamic EWS specification accounting both for endogenous and exogenous crisis persistence takes the form of:⁵

$$\Pr_{t-1}(y_t = 1) = F(\pi_t) = F(\gamma + \alpha y_{t-1} + x_{t-1}\beta + \delta \pi_{t-1}), \quad (1)$$

where $\Pr_{t-1}(y_t = 1)$ is the conditional probability of observing a currency crisis given the information set we have at our disposal at time $t - 1$, and π_t is the index at time t . F is the logistic c.d.f., which is preferred to the gaussian one as it is more appropriate for the study of extreme events such as crises. It is worth nothing that this method can be applied so as to construct EWSs for any type of crisis (banking, currency, sovereign debt, etc.) as long as the dependent variable is dichotomic.

The main advantage of this general framework is that it allows us to easily estimate and then compare different alternative EWS specifications under the same exact maximum likelihood (EML) framework. We first consider the pure static model, as in the rest of the literature in which the occurrence of currency crises is explained only by exogenous macroeconomic variables (x_{t-1}). This constitutes the benchmark model, in which crises are persistent only if the changes in economic indicators are themselves persistent (exogenous persistence). At the same time, we propose dynamic EWS specifications where we also allow for the endogenous persistence of a crisis. Formally, the dynamics can be introduced in two ways: either through the lagged binary variable y_{t-1} , or through the lagged index π_{t-1} .

The first dynamic model includes the lagged value of the binary dependent variable y_{t-1} . In this case, we can assess the impact of the regime prevailing in the previous period on the crisis probability. Note the existence of threshold effects, as the index must go beyond a certain threshold to set off a crisis in the previous period, *i.e.* $y_{t-1} = 1$. A second dynamic model includes the lagged index π_{t-1} . This time an increase in the index is linearly transmitted to the next period, hence always increasing the probability of

⁵It can be noted that the model for multi-step-ahead forecasts can be obtained by repetitive substitutions. For more details, see Kauppi and Saikkonen (2008). In addition, different lags of the exogenous variables can be considered, but in this paper we restrict our attention to only one lag ($t - 1$) as is generally done in the EWS literature.

crisis. Finally, the most complex dynamic model combines the two previous cases and includes both the lagged dependent variable y_{t-1} and the lagged index π_{t-1} . Nevertheless, one should not lose sight of the fact that since in the last two models δ is an autoregressive parameter it has to satisfy the usual stationarity condition, *i.e.* the roots of the corresponding polynomial lie outside the unit circle. Otherwise, the crisis becomes perpetual, which is counterintuitive. To tackle this problem, a constrained maximum likelihood estimation method is implemented and described in Appendix A. In fact, the log-likelihood function takes the general form of:

$$\text{LogL}(\theta) = \sum_{t=1}^T l_t(\theta) = \sum_{t=1}^T [y_t \log F(\pi_t(\theta)) + (1 - y_t) \log(1 - F(\pi_t(\theta)))], \quad (2)$$

where θ is the vector of parameters. The EML estimators have the desired large-sample properties. Furthermore, we tackle the autocorrelation problem induced by the construction of a j -month-ahead crisis variable by considering a Gallant correction for the covariance matrix. Given the maximum-likelihood framework, these dynamic time-series models are easy to implement using any existing econometric software.

2.2 Panel Framework

Instead of considering EWSs for individual countries (in a time-series approach), several papers (Berg and Patillo, 1998; Kumar et al., 2003 *inter alia*) favor a panel data framework (with fixed effects) by pooling the information available in several countries. In this section we thus propose a dynamic version of the fixed effects panel model by elaborating on Carro (2007). Note that this framework can embody all of the three dynamic specifications proposed in time-series. This constitutes an important contribution of our paper to the EWS literature.

Let us consider the panel version of the general dynamic binary-choice EWS:

$$\Pr(y_{it} = 1) = F(\eta_i + \alpha y_{i,t-1} + x_{i,t-1}\beta + \delta \pi_{i,t-1}), \text{ for } t = 0, 1, 2, \dots, T, \text{ and } i = 1, 2, \dots, N, \quad (3)$$

where N is the number of countries in the panel, T represents the number of time series observations for each country, and η_i accounts for the permanent unobserved heterogeneity between countries. Since we do not make any assumptions about the distribution of $\{\eta_i\}_{i=1}^N$, they are treated as parameters to be estimated, and our approach is a fixed effects one. In addition, we assume no cross-sectional dependence.

The log-likelihood of the model conditioned on the first observation, often called concentrated likelihood, takes the following form:

$$\text{LogL}(\theta, \eta_i) = \sum_{i=1}^N \text{LogL}_i(\theta, \eta_i) = \sum_{i=1}^N \sum_{t=1}^T [y_{it} \log(F_{it}) + (1 - y_{it}) \log(1 - F_{it})], \quad (4)$$

where θ represents the vector of parameters. As usual, the estimated parameters maximize the log-likelihood function, which means that they solve the first order conditions with respect to θ and with respect to η_i . Most importantly, the estimation of θ depends on $\hat{\eta}_i$, which means that $\hat{\theta}$ is a convergent estimator of θ_0 only when $\hat{\eta}_i$ is a convergent estimator of η_{i0} , that is when $T \rightarrow \infty$. Thus, the central issue here, as in any non-linear panel model with fixed effects, is how to deal with this incidental parameters problem.

The solution proposed by Carro (2007) actually consists in a numerical substitution of the fixed effects (η_i) in the estimation of θ . Thus, at each step N non-linear equations are solved to estimate $\{\hat{\eta}_i\}_{i=1}^N$ by using the $\hat{\theta}$ obtained in the previous step and then the estimated values of $\hat{\eta}_i$ are introduced into the first order condition corresponding to the concentrated likelihood so as to estimate $\hat{\theta}$. To be more precise, the

estimation of η is nested in the algorithm that maximizes the concentrated log-likelihood. At each iteration, $N + 1$ non-linear optimizations are realized using the Gauss-Newton algorithm (the first N ones correspond to the fixed effects, while the last one corresponds to the θ parameters). Carro's correction applied to our fixed-effects panel model is detailed in Appendix B. The main advantage of Carro's (2007) estimation method consists in its simplicity of implementation since it is based on the first derivatives of the log-likelihood function. Furthermore, it allows for the estimation of all the dynamic specifications introduced in the previous subsection. In particular, it allows one to proceed with a constrained estimation (as described in Appendix A) to assure the stationarity of the autoregressive parameter δ . Finally, given the reduction of the bias, the estimators have good asymptotic properties.

3 Empirical Application: Static vs Dynamic EWS

To prove the importance of crisis dynamics in the specification of an EWS, in this section we tackle the forecasting abilities of our dynamic EWSs relatively to those of the traditional static binary EWS. The dynamic EWS models we propose can be estimated for all the types of crises, *i.e.* banking, currency, debt, etc., as long as the dependent variable is dichotomic, but we restrict our attention here to currency crises. Nevertheless, before starting the analysis *per se*, let us present some data-related issues.

3.1 Dataset

Monthly data expressed in US dollars covering the period 1985-2011 for 16 emerging countries⁶ have been extracted from the IMF-IFS database and the national banks of the countries under analysis via Datastream. Several macroeconomic variables from two economic sectors were selected (see Candelon et al. 2012, Berg et al., 2008, Lestano et al., 2003):

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 domestic credit over GDP, the one-year growth rate of domestic credit over GDP, the real interest rate, and the real exchange rate overvaluation.

As in Kumar (2003), we treat outliers by dampening every variable using the formula: $f(x_t) = \text{sign}(x_t) * \ln(1 + |x_t|)$, so as to reduce the impact of extreme values. The series have been tested for non-stationarity using Augmented Dickey-Fuller tests. Additionally, the gaps through the series are replaced by cubic spline interpolation.

3.2 Dating Currency Crises

The most common method to identify currency crisis periods implies the computation of an index of speculative pressure. If this index exceeds a certain threshold, a crisis episode is identified. Note however that given the imperfections of existing crisis dating methods, the crisis periods identified in this way are cross-checked with the historical evidence of currency crises in each country. As in Candelon et al. (2012), we base our

⁶Argentina, Brazil, Chile, Colombia, Ecuador, Indonesia, Malaysia, Mexico, Peru, Philippines, South Africa, South Korea, Turkey, Thailand, Uruguay and Venezuela.

choice on the results of Lestano and Jacobs (2004).⁷ We hence identify crisis periods using the KLR modified pressure index (KLRm), which, unlike the traditional KLR index, also includes interest rates:

$$\text{KLRm}_{n,t} = \frac{\Delta e_{n,t}}{e_{n,t}} - \frac{\sigma_e}{\sigma_r} \frac{\Delta r_{n,t}}{r_{n,t}} + \frac{\sigma_e}{\sigma_i} \Delta i_{n,t}, \quad (5)$$

where $e_{n,t}$ denotes the exchange rate (*i.e.*, units of country n 's currency per US dollar in period t), $r_{n,t}$ represents the foreign reserves of country n in period t , while $i_{n,t}$ is the interest rate in country n at time t . Meanwhile, the standard deviations σ_X are actually the standard deviations of the relative changes in the variables $\sigma(\Delta X_{n,t}/X_{n,t})$, where X denotes each variable separately, including the exchange rate and the foreign reserves, with $\Delta X_{n,t} = X_{n,t} - X_{n,t-6}$. Furthermore, we use the standard deviations of the absolute changes in interest rates $\sigma_{\Delta i_{n,t}}$.⁸ The threshold equals two standard deviations above the mean, so that the crisis at time t indicator is given by:

$$C1_{n,t} = \begin{cases} 1, & \text{if } \text{KLRm}_{n,t} > 2\sigma_{\text{KLRm}_{n,t}} + \mu_{\text{KLRm}_{n,t}} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

From a macroeconomic point of view, it is more important to know if there will be a crisis within a certain horizon than in a certain month, because this time period allows the authorities to take steps to prevent the crisis. Consequently, we also gauge the forecasting abilities of our dynamic EWSs based on a $C6_{n,t}$ variable, which serves as the crisis dummy variable for each country, taking the value of 1 if there will be at least one crisis in the following 6 months and 0 otherwise:

$$C6_{n,t} = \begin{cases} 1, & \text{if } \sum_{j=1}^6 C1_{n,t+j} > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Note that hereafter the variables $C1_{n,t}$ and $C6_{n,t}$ (hereafter $C1$ and $C6$, respectively) correspond to y_t from our general framework.

3.3 Preliminary Analysis

The first step in our empirical analysis consists in selecting the explanatory variables from the list presented in subsection 3.1. For this, we estimate the traditional static logit model in a time-series framework. The whole dataset is considered (January 1986 - February 2011) and the dependent variable represents the occurrence of a crisis at time t , $C1$. The set of explanatory variables retained for the analysis includes: 1) the one-year growth rate of international reserves, 2) the one-year growth rate of imports, 3) the one-year growth rate of exports, 4) the one-year growth rate of M2 to foreign reserves, 5) the domestic credit over GDP and 6) the real interest rate. This selection is based on previous results found in the literature, on the correlation between the indicators, and on the explanatory power of each variable. Note that intuitively the calibration of this procedure on the static specification involves the static model going into the model comparison step

⁷KLR is often considered as too broad as a measure of currency crisis. For such a reason alternative approaches for dating crises have been proposed in the literature, See inter alii Zhang (2001). Nevertheless, KLRm remains the most commonly used one.

⁸Additionally, we take into account the existence of higher volatility in periods of high inflation, and consequently the sample is split into high and low inflation periods. The cut-off corresponds to a six-month inflation rate higher than 50%.

with a certain advantage relative to the dynamic models.

Next, we estimate our benchmark model, *i.e.* the static EWS model (labeled Model 1) as well as three types of dynamic EWSs: one including the *lagged binary* dependent variable (labeled Model 2), a dynamic one including the *lagged index* (called Model 3) and, last but not least, a dynamic model which includes both the *lagged binary* dependent variable and the *lagged index* (called Model 4).

Recall that the objective of this paper is to gauge the forecasting abilities of the dynamic EWSs relatively to those of the static one. Accordingly, the alternative specifications are not compared by relying on the estimated parameters, *i.e.* the signs, significance and economic interpretation of the estimators, but only based on their forecasting abilities.⁹ Note however that the coefficient of the lagged binary dependent variable is most of the time significant and has a positive sign. It follows that the probability of being in a crisis increases if a crisis regime prevailed in the previous period. This clearly indicates that crisis persistence should be taken into account in the construction of currency crisis EWS.

More formally, we select the most parsimonious dynamic specification by relying on the Schwarz Information Criterion (hereafter SBC). For comparison reasons, we also include the values of SBC for the static model. This goodness of fit indicator reveals that the right-hand-side variables have important explanatory power, especially when the lagged dependent variable and/or the lagged index are present in the model, *i.e.* the dynamic models (see Table 1).

insert Table 1

The lowest values of the SBC criterion correspond to the dynamic models and in particular to Model 2, *i.e.* including the lagged crisis indicator, which seems to be the most adequate dynamic specification for most of the countries (8 out of 16). To put it another way, the goodness of fit indicator provides evidence of the fact that in-sample dynamic specifications generally outperform the static one. The introduction of the lagged binary dependent variable into the model leads to the largest improvement in currency crisis EWSs. It is this dynamic specification that we retain for the subsequent forecasting exercise.

4 Forecast Evaluation

But are these new EWSs characterized by better forecasting abilities than the traditional static EWS? In this section we go one step further and statistically test the in-sample and out-of-sample forecasting abilities of the static and dynamic (including y_{t-1}) currency crisis EWS models. For this, we apply the EWS evaluation methodology developed in Candelon et al. (2012).¹⁰

4.1 In-Sample Analysis

As in the preliminary analysis, we check the in-sample forecasting abilities of the static and dynamic time-series models for the whole dataset (January 1986 - December 2011). First we consider the case where the event studied is *crisis at time t* , $C1$, and then we tackle the more insightful situation where the event under analysis is *crisis within 6 months*, $C6$. We assess the forecasting abilities of these alternative specifications by considering the *AUC* evaluation criterion as well as the *ROC* comparison test (Lambert and Lipkovich, 2008). As Candelon et al. (2012) argue, these tools are more suited to the evaluation of EWSs than classic measures like *QPS*, *LPS* and the tests associated with them.

⁹The estimated parameters are hence not reported but these results are available upon request from the authors.

¹⁰For more details see Appendix C.

The left-hand columns of Table 10 report the *AUC* performance assessment criterion and *ROC* test statistic and p-value for the two time-series models, *i.e.* static and dynamic logit, when using the *C1* dependent variable.

insert Table 10

Recall that the higher the *AUC*, the better the model. It is thus evident that the outperforming EWS for most of the countries (12 out of 16) is the dynamic logit. The results are similar for both Latin-American and Asian countries, the only cases where the static model seems to do better being Peru and South Korea.

To better grasp the differences in forecasting abilities between the two models we then rely on the *ROC* test. This a test allows one to check if the differences between the performance of the models are significant or not and for which countries. We identify 7 countries, Argentina, Brazil, Chile, Philippines, South Africa, Turkey and Uruguay, for which the dynamic time-series specification outperforms the static one and only one, South Korea, for which the results are reversed. In addition, the test provides statistical evidence that for the other countries, *e.g.* Colombia, Ecuador, Indonesia, etc., the static and dynamic models have similar forecasting abilities. Indeed, the *ROC* test rejects the null hypothesis of equal forecasting abilities for 8 out of the 16 countries. It thus corroborates our main finding that accounting for the endogenous persistence of the crisis by introducing the lagged binary dependent variable into the model matters for the forecasting abilities of currency crisis EWSs.

Given that the panel framework has received a lot of attention in the EWS literature, we now compare the forecasting abilities of time-series versus panel EWS models. We consider the dynamic time-series and panel models in the form including the *lagged binary* dependent variable along with the selected macroeconomic indicators. Both a fixed-effects panel which uses the whole sample of countries, and a fixed-effects panel model based on two regional clusters (for Latin - America and Asia¹¹) are estimated.

The results displayed in Table 3 provide clear evidence in favor of the time-series specification. The *AUC* evaluation criterion (see the left-hand columns of Table 3) is larger for the time-series logit model in 13 out of the 16 countries in the sample. The results of the *ROC* comparison test comparing the time series and the panel dynamic models are presented in the right-hand columns of Table 3. The null hypothesis of similar forecasting abilities cannot be rejected for either of the panel models in 13 out of the 16 countries considered. And when it is rejected, as for Chile, Ecuador and Turkey, it provides evidence in favor of the dynamic country-by-country model. Consequently, the dynamic panel models do not have better forecasting abilities than the dynamic time series EWS, thus supporting the results obtained by Berg et al (2008).

To summarize, our findings indicate that there are gains from using a dynamic EWS specification which includes the lagged binary crisis indicator.¹²

We next check if these findings still hold when the forecast horizon is longer. Intuitively, this larger horizon is compatible with a global perspective on the state of the economy in the following periods. This is a fundamental requirement for an EWS designed to give regulators enough time to react appropriately to reduce the likelihood of a crisis. We decide to consider a 6-month horizon.¹³ The event under analysis, denoted by *C6*, corresponds to the occurrence of a crisis in the following six months (see section 3.2 for the construction of this variable).

¹¹South Africa is included in the Asian cluster.

¹²We also performed a robustness check of our results by considering the *Zhang* pressure index instead of the *KLRm* for the identification of calm and crisis periods. Similar results have been obtained and they are available upon request.

¹³We thank an anonymous referee for suggesting this forecasting horizon. The output obtained with a forecast horizon of 24 months (the horizon considered in Kaminski et al., 1998 and Berg and Patillo, 1999) are qualitatively similar and available in an earlier version of the paper or upon request from the authors.

Our predictive model aims to estimate the probability in $t+1$ of the occurrence of a crisis in the following 6 periods, $\Pr(C6_{t+1} = 1)$, where the current period is denoted by t . Consequently, it must be built only on information available at time t . In a real-time exercise the existing information regarding the occurrence of crises in the last 6-month period, $C1_{t-5} : C1_t$ is synthesized by the 6th lag of the $C6$ variable, $C6_{t-6}$, (see Figure 1). Therefore, the probability of having at least one crisis in the following 6 months, $\Pr(C6_{t+1} = 1)$, should depend on the exogenous macroeconomic variables, x_t , and on the state of the economy during the previous 6 months, $C6_{t-6}$. To be more precise, the model we estimate takes the form:

$$\Pr_t(C6_{t+1} = 1) = F(\alpha + \delta C6_{t-6} + x_t \beta). \quad (8)$$

Figure 1: Forecasting a crisis within 6 months

The results for the evaluation of the forecasting abilities of the static and the dynamic logit models using $C6$ as dependent variable go along the same lines as those obtained when using $C1$ on the left-hand-side of the regression (see the right-hand part of Table 10). The *AUC* criterion associated with the dynamic model is larger than that of the static EWS in 12 of the countries. This finding suggests that the introduction of the crisis dynamics in the form of the lagged crisis indicator into the EWS specification improves the model's forecasting abilities not only one-period ahead but also within a six-month interval. The *ROC* test supports the previous result by identifying 5 countries, *i.e.* Argentina, Colombia, Philippines, South Korea and Turkey, for which the dynamic logit model performs significantly better in forecasting the occurrence of a crisis in the following 6 months than the static one. Note also that the static model does not outperform the dynamic one in any case; the two models have equal forecasting abilities for all the other countries in the sample.

Nevertheless, the best model in-sample is not necessarily the outperforming one out-of-sample. The same statistical assessment framework is hence applied in the following section for the out-of-sample forecasting performance of static and dynamic models.

4.2 Out-of-sample analysis

To check the out-of-sample performance of our static and dynamic time-series EWS models we consider a rolling forecasting scheme. The sample is split at the end of 1996, to allow for most of the countries, *i.e.* 14, to experience at least one crisis in the out-of-sample period. On one hand, for the crisis at time t analysis, this implies that the dynamic time-series logit model is estimated over the January 1986 - December 1996

window and the estimated parameters are used to compute the probability of there being a crisis in January 1997. These estimation and forecasting steps are then repeated for the February 1986 - January 1997 period to obtain the out-of-sample crisis probability for February 1997 and so on.

On the other hand, for the crisis in the following 6 months analysis, the model is estimated over the January 1986 - December 1996 window by equation (8) and the probability of there being a crisis in the 6 months starting from January 1997 is computed. The other elements of the sequence of crisis probabilities in 6 months are constructed in the same way, just by shifting the estimation window to the right by one period. Note that the out-of-sample forecast series are evaluated by using the same methodology of Candelon et al. (2012) as previously.

The left-hand columns of Table 4 report the *AUC* and *ROC* test statistic and p-value for the out-of-sample forecasts issued by the static and the dynamic time-series logit models for a one-month horizon. For the crisis in 6 months analysis the same evaluation measures are displayed in the right-hand part of the same table.

insert Table 4

It appears that the dynamic logit model outperforms the static one for some of the countries, both when considering the crisis at time t and crisis within 6 months events. This finding reinforces the in-sample result that crisis dynamics should be taken into account to improve the forecasting abilities of an EWS. To be more precise, according to the *AUC* criterion, the dynamic model fares better than the static one in half of the countries. Most importantly, the differences in forecasting abilities are significant and in favor of the dynamic specification for 2 (resp. 5) countries in the case of *C1* (resp. *C6*). For all the others (except Thailand - in the first analysis - and Venezuela - in the second analysis) the two models perform equally well.

Furthermore, Figures (2) and (11) present the out-of-sample crisis at time t probabilities output by the static and dynamic logit models for each of the countries in the sample from January 1997 to December 2011. The optimal cut-off for the static model is also displayed, so that a crisis period is identified by this model if the associated probability goes beyond that threshold.¹⁴ Note that the optimal cut-off is the one that best discriminates between crisis and calm periods. It is identified by using the in-sample dataset (see Appendix C for more details).

insert Figures 2 and 11

The crisis probability forecast by the two alternative models seems very low in most of the calm periods while it is very high during crisis periods. The dashed line, which corresponds to the dynamic model, is nevertheless much closer to 1 during crises than the one for the static model in several countries (*e.g.* Brazil, Chile, Colombia, etc.). At the same time, during calm times, the sequence of probabilities obtained from the dynamic model is as least as low as the one issued by the static one.

For the crisis in 6 months event similar illustrations of the evolution of the out-of-sample probabilities are displayed in Figures (4) and (5).

inset Figures 4 and 5

¹⁴To keep the figures legible and unambiguous, we do not report both cut-offs (for the static and dynamic models).

As expected, the series of probabilities are slightly noisier, since we have increased the forecasting window from one to 6 months. The crises are generally correctly forecast, but there are also some false alarms, mainly around the global financial crisis. This provides evidence for the sensitivity of the exchange market to stress in the stock and sovereign ones. The stress is not intense enough to set off a currency crisis (according to the KLRm dating method), but the leading indicators clearly perceive this vulnerability and lead to high probabilities of crisis (see for example Argentina, Brazil, Philippines). A visual inspection of the crisis probabilities also suggests that the dynamic model better forecasts crisis and calm periods relative to the static one in countries like Brazil, Chile, Malaysia, Turkey, etc. As in the case of *C1*, the countries which never faced a crisis after 2006, *i.e.* Mexico and Peru, are characterized by very low crisis probabilities.

The last step in the out-of-sample analysis of the forecasting abilities of the two models consists in an analysis of the performance in absolute terms of the best EWS model for each country. The optimal cut-off, the percentage of correct predictions of crisis and calm periods as well as the corresponding EWS are reported in Table 5 for the two events studied (*C1* and *C6*). It seems that both the static and dynamic country per country models are characterised by small values of the cut-offs (they range between 0.003 and 0.375). These thresholds are generally larger for the crisis in 6 months event, in accordance with the larger variation in probabilities registered in this case.

In most of the countries, 10 (resp. 9) out of 14 for the *C1* (resp. *C6*) event, the model best identifying crisis and calm periods is the dynamic one. This means that the lagged dependent variable has improved explanatory power and discriminates very well between calm and crisis periods, which further motivates the use of dynamic EWSs. Moreover, both crisis and calm periods are very well forecast by the optimal model, *i.e.* the percentage of correct predictions lays between 44.4% and 100% for *C1* while in the case of *C6* it ranges between 42.9% and 100%. The relative forecasting abilities of the dynamic model remain the same if not improved when moving to a longer forecast horizon (*C6*), which is more desirable from the policymakers' perspective.

To sum up, the dynamic EWS model including the lagged binary crisis indicator keeps the best forecasting abilities not only in-sample but also out-of-sample.

4.3 Robustness Check using alternative currency crises indicators

In the previous sections of the paper, currency crises were dated using the KLRm approach described from eq. 5. Nevertheless, literature that has emerged subsequent to the global financial crisis, is offering a multitude of alternative currency crisis indicators.¹⁵ To assess that our main result (the dynamic EWS performs the static one) is a general fact and not specific to the KLRm indicator, the in- and out-of-sample analysis is repeated considering alternative currency crisis indicators.

Two categories of crisis dating methods are generally considered in the literature. The first one relies on exchange market pressure indices whereas the second one is based on the absolute values of macroeconomic variables.

The KLRm indicator belongs to the first category. To complete it, we also consider the Zhang method, which considers time-varying thresholds (equal to 3 and -3, respectively) both for the exchange rate and the international reserves component.

Within the second category, we consider two alternative crisis indicators. The one proposed by (Frankel and Saravelos, 2012; Laeven and Valencia, 2012) which relies on nominal exchange rates depreciation. A crisis is detected when depreciation exceeds 10% on a year to year basis. The second one is developed by

¹⁵See Frankel and Saravelos (2012) for a recent description of existing indicators.

Ghosh and Ghosh (2003) and analyses the real impact of a currency crisis. It indicates a crisis when real GDP growth falls by more than 30%.

The outcomes of the obtained by considering each of these four alternative definitions of a currency crisis are displayed in tables The *ROC* test indicates that the dynamic EWS specification performs better than the static one in most of the countries regardless of the way the crisis indicator is constructed.¹⁶ This finding holds not only for the crisis at time t event but also for the crisis within 6 months horizon. It is interesting to notice that the superiority of the dynamic EWS model versus the static one is even more evident when considering currency crisis indicators based on the absolute values of the macroeconomic variables. The differences in the *ROC* test-statistics are significant in almost all countries (resp. more than half of the countries) contrary to about 6 (resp. 2) of them for the crisis at time t (resp. crisis within 6 months) event. These results hold not only in-sample but also out-of-sample (see table). This robustness check reinforces previous conclusions, emphasizing hence the importance of dynamic EWS for forecasting currency crises.

5 Conclusion

This paper provides evidence of the importance of taking crisis dynamics into account in order to adequately forecast crises. It is the first paper to use Kauppi and Saikkonen’s exact ML methodology (2008) to estimate a dynamic discrete-choice EWS. Furthermore, it extends this methodology to panel by drawing on work by Carro (2007). The forecasting abilities of static and dynamic binary EWSs are then compared and discussed in a unified evaluation framework specifically designed for EWS (Candelon et al. 2012).

Several conclusions can be drawn from the empirical application of this methodology to construct currency crisis EWS models. First, we show that dynamic logit models consistently outperform the static one in the in-sample forecast exercise. This result is corroborated by the out-of-sample analysis performed by considering as the dependent variable both the crisis at time t and the crisis in 6 months indicators. In particular, it appears that the forecasting abilities of a dynamic model improve, or at least stay the same when increasing the forecast horizon ($C6$). In addition, the optimal EWS for each country yields good forecasting probabilities, which are generally beyond the optimal cut-off in the crisis periods and below this threshold the rest of the time.

There is no doubt that in the quest for a new generation of financial crisis EWSs, dynamics should constitute a key characteristic that would help deliver more adequate signals of impending financial turmoils. Let us hope that policy makers can exploit these signals to tame such painful events.

Bibliography

1. Abiad, A., (2003). Early Warning systems: A Survey and a Regime Switching Approach, IMF Working Paper.
2. Arias, G., and Erlandsson, G., (2005). Improving Early Warning Systems with a Markov Switching Model - an Application to South-East Asian Crises, Lund University, Working Paper.
3. Bai, J., and Ng, S., (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.

¹⁶The *AUC* evaluation criterion leads to the same conclusion. For the sake of space we do not report these results, but they are available from the authors upon request.

4. Basel Committee on Banking Supervision, (2005). Studies on the Validation of Internal Rating Systems, working paper no.14, Bank for International Settlements.
5. Berg, A., and Pattillo, C., (1999). Predicting Currency Crises: The Indicators Approach and an Alternative. *Journal of International Money and Finance*, 18, 561-586.
6. Berg, J., Candelon, B., and Urbain, J.P., (2008). A Cautious Note on the Use of Panel Models to Predict Financial Crises, *Economics Letters* 101, issue 1, 80-83.
7. Berg, A., and Cooke, R., 2004, Autocorrelation Corrected Standard Errors in Panel Probits: an Application to Currency Crisis Prediction, IMF Working Paper 39.
8. Bussiere, M., and Fratzscher, M., (2006). Towards a New Early Warning System of Financial Crises, *Journal of International Money and Finance*, 25(6), 953-973.
9. Candelon, B., Dumitrescu, E.I., and Hurlin, C., (2009). Towards a Unified Framework to Evaluate Financial Crisis Early Warning Systems, METEOR Research memorandum RM/10/046.
10. Candelon, B., Dumitrescu, E.I., Hurlin, C., (2012). How to evaluate an Early Warning System? Towards a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods, *IMF Economic Review*, 60(1), 75-113.
11. Carro, J.M., (2007). Estimating dynamic panel data discrete choice models with fixed effects, *Journal of Econometrics*, 140, 503-528.
12. Clark, T. E., West, K. D., (2007). Approximately Normal Tests for Equal Predictive Accuracy in Nested Models, *Journal of Econometrics*, 138(1), 291-311.
13. Dabrowski, M. and Jakubiak, M., (2003). The Sources of Economic Growth in Ukraine after 1998 Currency Crisis and the Country's Prospects, CASE Network Reports 0055, CASE-Center for Social and Economic Research.
14. Diebold, F. X., Mariano, S., (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, American Statistical Association, 13(3), 253-63.
15. Falcetti, E., Tudela, M., (2006). Modelling Currency Crises in Emerging Markets: A Dynamic Probit Model with Unobserved Heterogeneity and Autocorrelated Errors, *Oxford Bulletin of Economics and Statistics*, 68(4), pages 445-471.
16. Gallant, A. R., (1987). *Nonlinear Statistical Models*, John Wiley and Sons, New York.
17. Frankel, J. and Saravelos, G., (2010) Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability ? Evidence from the 2008-09 Global Crisis, NBER Working Papers 16047.
18. Fuertes, A.-M., Kalotychou, E., (2007) Optimal Design of Early Warning Systems for Sovereign Debt Crises, *International Journal of Forecasting*, 23(1), 85-100.
19. Ghosh, W.R., Ghosh, A.R., (2003) Structural Vulnerabilities and Currency Crises, *IMF Staff Papers*, 50(3), 481-507.
20. Glick, R., and Hutchison, M., (1999). Banking and Currency Crises: How Common Are The Twins?, Working Papers 012000, Hong Kong Institute for Monetary Research.
21. Harding, D., and Pagan, A., (2011). An Econometric Analysis of Some Models for Constructed Binary Time Series, *Journal of Business & Economic Statistics*, vol. 29(1), 86-95.

22. Kaminsky, G., Lizondo, S., Reinhart, C., (1998). Leading Indicators of Currency Crises, IMF Staff Papers, 45(1), 1-48.
23. Kauppi, H., Saikkonen, P., (2008), Predicting U.S. Recessions with Dynamic Binary Response Models, The Review of Economics and Statistics, 90(4), 777-791.
24. Kumar, M., Moorthy, U., and Perraudin, W., (2003). Predicting Emerging Market Currency Crashes, Journal of Empirical Finance, 10, 427-454.
25. Lambert, J. and Lipkovich, I., (2008). A Macro for Getting More out of your ROC Curve, SAS Global forum, paper 231.
26. Lestano and Jacobs, J., (2004). A Comparison of Currency Crisis Dating Methods: East Asia 1970-2002, CCSO Working Papers 200412, CCSO Centre for Economic Research.
27. Martinez Peria, M.S., (2002). A regime-switching approach to the study of speculative attacks: A focus on EMS crises, Empirical Economics, 27(2), 299-334.
28. Rose, A.K. and Spiegel, M.M., (2010). Cross-Country Causes and Consequences of the 2008 Crisis: Early Warning, forthcoming in Global Journal of Economics.
29. Tudela, M., (2004). Explaining Currency Crises: a Duration Model Approach, Journal of International Money and Finance, 23(5), 799-816.
30. Zhang, Z. (2001), Speculative Attacks in the Asian Crisis, IMF Working Paper 189, International Monetary Fund, Washington, D.C.

Appendix A: Constrained Maximum Likelihood Estimation (Kauppi and Saikkonen, 2008)

Let us recall the general form of the model in the case of a logistic distribution function Λ : $\Pr_{t-1}(y_t = 1) = \Lambda(\gamma + \delta\pi_{t-1} + \alpha y_{t-1} + x_{t-1}\beta)$. Following Kauppi and Saikkonen, we set the initial value π_0 to $(\gamma + \bar{x}\beta)/(1 - \delta)$, where $\bar{x}$ is the sample mean of the exogenous variables. The initial condition for the constant, γ , and β parameters is given by an OLS estimation, while the initial δ is set to 0. Moreover, since δ is an autoregressive parameter, a constrained maximum likelihood estimation must be implemented. Nevertheless, the same results can be reached in a faster and easier way, by using a transformation of the δ parameter in the classical maximum likelihood process. To solve this problem, we denote by ψ the new maximization parameter, identified so that δ is equal to $\psi/(1 + |\psi|)$, *i.e.* δ takes values in the interval $[-1, 1]$.

The log-likelihood function takes the form of:

$$\text{LogL}(\theta) = \sum_{t=1}^T l_t(\theta) = \sum_{t=1}^T [y_t \log \Lambda(\pi_t(\theta)) + (1 - y_t) \log(1 - \Lambda(\pi_t(\theta)))], \quad (9)$$

where θ is the vector of parameters $\theta = [\gamma, \psi, \alpha, \beta]$.

Notice that in view of the parameter transformation from δ to ψ , the covariance matrix corresponds to the parameters $[\gamma, \psi, \alpha, \beta]$, and not to the initial parameters $[\gamma, \delta, \alpha, \beta]$. Subsequently, we must proceed with a change of the covariance matrix from the first parameter space to the second one. To this end, we use

Taylor's theorem to calculate the approximation of the transformation function around the point ψ_0 . Since the estimated parameter $\hat{\delta}$ is equal to $f(\hat{\psi})$, where $f(\hat{\psi}) = \hat{\psi}/(1 + |\hat{\psi}|)$, the approximation becomes:

$$\hat{\delta} = f(\hat{\psi}) \simeq f(\psi_0) + \left. \frac{\partial f(\hat{\psi})'}{\partial \psi} \right|_{\psi_0} (\hat{\psi} - \psi_0). \quad (10)$$

Nevertheless, we aim to find the variance of δ , and thus, using the formula $\mathbb{V}(a'X) = a'\mathbb{V}(X)a$, we obtain:

$$\mathbb{V}(\hat{\delta}) \simeq \left. \frac{\partial f(\hat{\psi})'}{\partial \psi} \right|_{\psi_0} \mathbb{V}(\hat{\psi}) \left. \frac{\partial f(\hat{\psi})}{\partial \psi} \right|_{\psi_0}. \quad (11)$$

Since $\hat{\psi} \xrightarrow{P} \psi_0$, we can replace ψ_0 with the estimator $\hat{\psi}$ in eq. (11)

$$\mathbb{V}(\hat{\delta}) \simeq \left. \frac{\partial f(\hat{\psi})'}{\partial \psi} \right|_{\hat{\psi}} \mathbb{V}(\hat{\psi}) \left. \frac{\partial f(\hat{\psi})}{\partial \psi} \right|_{\hat{\psi}}, \quad (12)$$

where the first derivative of the transformation function $f(\hat{\psi})$ with respect to $(\hat{\psi})$ can be computed through finite differences. Consequently, the standard errors obtained as the square root of the elements lying on the first diagonal of the covariance matrix are consistent with the $[\gamma, \delta, \alpha, \beta]$ vector of parameters. More exactly, a Gallant correction based on a Parzen kernel (Gallant, 1987) is used for the covariance matrix. Kauppi and Saikkonen (2008) argue that robust standard errors can be obtained as the diagonal elements of the matrix $\hat{J}(\hat{\theta})^{-1} \hat{I}(\hat{\theta}) \hat{J}(\hat{\theta})^{-1}$, where $\hat{I}(\hat{\theta}) = T^{-1}(\sum_{t=1}^T \hat{d}_t' \hat{d}_t + \sum_{t=1}^T w_{Tj} \sum_{t=j+1}^T (\hat{d}_t' \hat{d}_{t-j} + \hat{d}_{t-j}' \hat{d}_t))$, $\hat{d}_t = \partial l_t(\hat{\theta}) / \partial \theta$ is the score for the t^{th} observation in the sample, and where $J(\theta) = \text{plim}_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T (\partial^2 l_t(\theta) / \partial \theta \partial \theta')$ is the asymptotic limit of the hessian. We consider that the robust covariance matrix should be used not only for multi-periods-ahead forecasts, (as in Kauppi and Saikkonen, 2008) but also for one- period-ahead forecasts for two reasons. First, the logistic distributional hypothesis imposed to the error term might not always hold and second, this corrected covariance matrix is robust to autocorrelation, automatically introduced when considering an EWS (see Berg and Coke, 2004).¹⁷

¹⁷Harding and Pagan, (2011) note that the binary variables dating crises, recessions and bulls and bears are 'constructed binary variables' which, contrary to those encountered in microeconomics, are not i.i.d. When constructed binary variables are used as a dependent variable, test statistics robust to the serial correlation and heteroskedasticity characterizing the binary variable must be employed.

Appendix B: Modified Maximum Likelihood Estimation (Carro, 2007)

As previously mentioned, the dynamic panel logit model with fixed effects is estimated by solving $N + 1$ non-linear equations based on the modified score of each country i , which takes the following form:

$$d_{\theta M_i}(\theta) = d_{\theta C_i}(\theta, \hat{\eta}_i(\theta)) - \frac{1}{2} \frac{1}{d_{\eta\eta i}(\theta, \hat{\eta}_i(\theta))} \left(d_{\theta\eta\eta i}(\theta, \hat{\eta}_i(\theta)) + d_{\eta\eta\eta i}(\theta, \hat{\eta}_i(\theta)) \frac{\partial \hat{\eta}_i(\theta)}{\partial \theta} \right) + \frac{\partial/\partial \eta_i(\mathbb{E}[d_{\theta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i])}{\mathbb{E}[d_{\eta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i]} \Big|_{\eta_i=\hat{\eta}_i(\theta)} \frac{\mathbb{E}[d_{\theta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i]}{\mathbb{E}[d_{\eta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i]} \Big|_{\eta_i=\hat{\eta}_i(\theta)} * \frac{\partial/\partial \eta_i(\mathbb{E}[d_{\eta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i])}{\mathbb{E}[d_{\eta\eta_i}(\theta, \eta_i)|y_{i0}, \eta_i, x_i]} \Big|_{\eta_i=\hat{\eta}_i(\theta)}, \quad (13)$$

where $d_{\theta C_i}(\theta, \hat{\eta}_i(\theta))$ is country i 's score from the concentrated likelihood (MLE):

$$d_{\theta C_i}(\theta, \hat{\eta}_i(\theta)) = \frac{y_{it} - F_{it}(\theta, \eta_i)}{F_{it}(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i))} \left(y_{it} + \frac{\partial \hat{\eta}_i(\theta)}{\partial \theta} \right), \quad (14)$$

θ is the vector of parameters (except for the country fixed effects η_i , $i = \{1, 2, \dots, N\}$ and $F(\cdot)$ is the logistic c.d.f.

From the first order condition of η_i , $d_{\eta_i}(\theta, \eta_i) = \sum_{t=1}^T (y_{it} - F_{it}(\theta, \eta_i))/F_{it}(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i)) = 0$, it can be derived that the estimators $\hat{\eta}_i$, $i = 1, 2, \dots, N$ solve the following equation:

$$\sum_{t=1}^T y_{it} \frac{f_{it}(\theta, \eta_i)}{F_{it}(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i))} = \sum_{t=1}^T \frac{F_{it}(\theta, \eta_i) f_{it}(\theta, \eta_i)}{F_{it}(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i))}. \quad (15)$$

Deriving eq. (14) with respect to θ for the second dynamic binary specification, *i.e.* the one including the lagged binary indicator, we can obtain $\partial \hat{\eta}_i(\theta)/\partial \theta$:

$$\partial \hat{\eta}_i(\theta)/\partial \theta = - \frac{\sum_{t=1}^T \frac{X_i Z}{F_{it}(\theta, \eta_i)^2 (1 - F_{it}(\theta, \eta_i)^2)}}{\sum_{t=1}^T \frac{Z}{F_{it}(\theta, \eta_i)^2 (1 - F_{it}(\theta, \eta_i)^2)}}, \quad (16)$$

where $X_i = y_{t-1}, x_{i1}, x_{i2}, \dots, x_{iK}$ is the explanatory variable corresponding to the $\theta = \alpha, \beta_1, \beta_2, \dots, \beta_K$ parameter we analyze, K is the number of explanatory variables, and $Z = y_t[f'_{it}(\theta, \eta_i)F_{it}(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i)) - f_{it}^2(\theta, \eta_i)(1 - 2F_{it}(\theta, \eta_i))] - f_{it}^2(\theta, \eta_i)F_{it}^2(\theta, \eta_i) - F_{it}^2(\theta, \eta_i)(1 - F_{it}(\theta, \eta_i))f'_{it}(\theta, \eta_i)$. Since $F_{it}(\theta, \eta_i)$ is the c.d.f., $f_{it}(\theta, \eta_i)$ is its associated density function and $f'_{it}(\theta, \eta_i)$ is the first derivative of the density function. For the logistic distribution $F_{it}(\theta, \eta_i) = \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i)/(1 + \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i))$, $f_{it}(\theta, \eta_i) = \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i)/(1 + \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i))^2$, and $f'_{it}(\theta, \eta_i) = \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i)(1 - \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i))/(1 + \exp(\alpha y_{it-1} + x_{it-1}\beta + \eta_i))^3$.

The partial derivative of the η gradient with respect to θ is given by the implicit functions theorem:

$$\partial \hat{\eta}_i(\theta)/\partial \theta_k = - \frac{\partial d_{\eta_i}(\theta, \hat{\eta}_i)/\partial \theta_k}{\partial d_{\eta_i}(\theta, \hat{\eta}_i)/\partial \eta_i}, \quad (17)$$

where $k=1,2,\dots,K$, K being the number of explanatory variables considered in the model, $\partial d_{\eta_i}(\theta, \hat{\eta}_i)/\partial \theta_k =$

$\partial^2 \text{LogL}(\theta, \eta_i) / \partial \eta_i \partial \theta_k |_{(\theta, \hat{\eta}_i)}$, and $\partial d_{\eta_i}(\theta, \hat{\eta}_i) / \partial \eta_i = \partial^2 \text{LogL}(\theta, \eta_i) / \partial^2 \eta_i |_{(\theta, \hat{\eta}_i)}$. The estimation of the parameters by standard MLE is straightforward since $d_{\theta C_i}(\theta, \hat{\eta}_i(\theta)) = 0$ and $d_{\eta_i}(\theta, \eta_i) = 0$ can be easily computed and solved. However, aiming to reduce the estimation bias, the implementation of MMLE becomes compulsory. For this, further information regarding the expected value of the first order condition and the derivatives of this expectation is required.

In view of the MMLE estimation, we derive the following elements for the α parameter, corresponding to the lagged binary variable:

$$d_{\alpha \eta_i}(\theta, \eta_i) = \frac{\partial^2 \text{LogL}_i}{\partial \alpha \partial \eta_i} = - \sum_{t=1}^T y_{i,t-1} f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i), \quad (18)$$

$$d_{\eta_i \eta_i}(\theta, \eta_i) = \frac{\partial^2 \text{LogL}_i}{\partial^2 \eta_i} = - \sum_{t=1}^T f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i), \quad (19)$$

$$d_{\alpha \eta_i \eta_i}(\theta, \eta_i) = - \sum_{t=1}^T y_{i,t-1} f'_{it}(\alpha y_{it-1} + \beta x_i + \eta_i), \quad (20)$$

$$d_{\eta_i \eta_i \eta_i}(\theta, \eta_i) = - \sum_{t=1}^T f'_{it}(\alpha y_{it-1} + \beta x_i + \eta_i), \quad (21)$$

where f is the logistic pdf and f' is its first derivative.

Next, we obtain the expectation of the derivatives $d_{\alpha \eta_i}$ and $d_{\eta_i \eta_i}$. In a first step we calculate $\Pr(y_{it} = 1 | y_{i0}, \eta_i, x_i)$, *i.e.* the probability that a crisis will occur in country i at time t given the initial value of the binary dependent variable y_{i0} , the fixed effects η_i and the explanatory variables x_i . The starting point is set to $\Pr(y_{i1} = 1 | y_{i0}, \eta_i, x_i) = F_{it}(\alpha y_{i0} + \beta x_i + \eta_i)$. For $t > 1$,

$$\begin{aligned} \Pr(y_{it} = 1 | y_{i0}, \eta_i, x_i) &= \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) (F_{it}(\alpha + \beta x_i + \eta_i) - F_{it}(\beta x_i + \eta_i)) \\ &\quad + F_{it}(\beta x_i + \eta_i). \end{aligned} \quad (22)$$

Note also that $\Pr(y_{i0} = 1 | y_{i0}, \eta_i, x_i) = y_{i0}$.

In the second step the expectation of the two derivatives can be calculated:

$$\mathbb{E}[d_{\alpha \eta_i}(\theta, \eta_i) | y_{i0}, \eta_i, x_i] = - \sum_{t=1}^T \mathbb{E}[y_{it-1} f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i) | y_{i0}, \eta_i, x_i], \text{ where} \quad (23)$$

$$\mathbb{E}[y_{it-1} f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i) | y_{i0}, \eta_i, x_i] = f_{it}(\alpha + \beta x_i + \eta_i) \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i). \quad (24)$$

and

$$\mathbb{E}[d_{\eta_i \eta_i}(\theta, \eta_i) | y_{i0}, \eta_i, x_i] = - \sum_{t=1}^T \mathbb{E}[f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i) | y_{i0}, \eta_i, x_i], \text{ where} \quad (25)$$

$$\begin{aligned} \mathbb{E}[f_{it}(\alpha y_{it-1} + \beta x_i + \eta_i) | y_{i0}, \eta_i, x_i] &= f_{it}(\alpha + \beta x_i + \eta_i) \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) \\ &\quad + f_{it}(\beta x_i + \eta_i) (1 - \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i)) \\ &= \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) (f_{it}(\alpha + \beta x_i + \eta_i) \\ &\quad - f_{it}(\beta x_i + \eta_i)) + f_{it}(\beta x_i + \eta_i). \end{aligned} \quad (26)$$

The last elements needed in the gradient function are the derivatives of the two expectations of $d_{\alpha\eta_i}$ and $d_{\eta_i\eta_i}$ with respect to the fixed effect η_i . Nevertheless, to compute these elements the derivative of the probability of occurrence of a crisis in country i at time t with respect to η_i must be calculated first:

$$\frac{\partial}{\partial \eta_i} \Pr(y_{i1} = 1 | y_{i0}, \eta_i, x_i) = f_{i1}(\alpha y_{i0} + \beta x_i + \eta_i) \text{ as starting point } (t=1) \text{ and} \quad (27)$$

$$\begin{aligned} \frac{\partial}{\partial \eta_i} \Pr(y_{it} = 1 | y_{i0}, \eta_i, x_i) &= \frac{\partial}{\partial \eta_i} \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) (F_{it}(\alpha + \beta x_i + \eta_i) \\ &\quad - F_{it}(\beta x_i + \eta_i)) \\ &\quad + \Pr(y_{it} = 1 | y_{i0}, \eta_i, x_i) (f_{it}(\alpha + \beta x_i + \eta_i) - f_{it}(\beta x_i + \eta_i)) \\ &\quad + f_{it}(\beta x_i + \eta_i). \end{aligned} \quad (28)$$

Finally, the last elements needed in the formula of the gradient are obtained:

$$\begin{aligned} \frac{\partial}{\partial \eta_i} (\mathbb{E}[d_{\alpha\eta_i}(\theta, \eta_i) | y_{i0}, \eta_i, x_i]) &= - \sum_{t=1}^T \frac{\partial}{\partial \eta_i} \mathbb{E}[y_{it-1} f(\alpha y_{it-1} + \beta x_{it-1} + \eta_i) | y_{i0}, \eta_i, x_i] \\ &= f'(\alpha + \beta x_{it-1} + \eta_i) \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) \\ &\quad + f(\alpha + \beta x_{it-1} + \eta_i) \frac{\partial}{\partial \eta_i} \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i), \end{aligned} \quad (29)$$

and respectively

$$\begin{aligned} \frac{\partial}{\partial \eta_i} (\mathbb{E}[d_{\eta_i\eta_i}(\theta, \eta_i) | y_{i0}, \eta_i, x_i]) &= - \sum_{t=1}^T \frac{\partial}{\partial \eta_i} \mathbb{E}[f(\alpha y_{it-1} + \beta x_{it-1} + \eta_i) | y_{i0}, \eta_i, x_i] \\ &= \frac{\partial}{\partial \eta_i} \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) (f(\alpha + \beta x_{it-1} + \eta_i) \\ &\quad - f(\beta x_{it-1} + \eta_i)) \\ &\quad + \Pr(y_{it-1} = 1 | y_{i0}, \eta_i, x_i) (f'(\alpha + \beta x_{it-1} + \eta_i) \\ &\quad - f'(\beta x_{it-1} + \eta_i)) + f'(\beta x_{it-1} + \eta_i). \end{aligned} \quad (30)$$

The individual score corresponding to the slope parameters, *i.e.* β , can be computed in a similar way. Nevertheless, we must take into account the fact that $\partial F / \partial \beta_k = x_{it,k} f$, where $x_{it,k}$ is the k^{th} explanatory variable, which is known and observable, contrary to y_{t-1} from the score corresponding to the lagged binary variable parameter α .

As for the covariance matrix, it is the inverse of the the MMLE Hessian matrix, which is calculated accounting for the fixed effects by using the following formula:¹⁸

$$\begin{aligned} \sum_{i=1}^N \left(\frac{\partial^2 \text{Log} L_i(\theta, \hat{\eta}_i(\theta))}{\partial \theta \partial \theta} + \frac{\partial^2 \text{Log} L(\theta, \eta_i)}{\partial \theta \partial \eta_i} \right) \Big|_{\eta_i = \hat{\eta}_i(\theta)} \frac{\partial \hat{\eta}_i(\theta)}{\partial \theta} \\ + \left[\frac{\partial^2 \text{Log} L_i(\theta, \eta_i(\theta))}{\partial \eta_i \partial \theta} \Big|_{\eta_i = \hat{\eta}_i(\theta)} + \frac{\partial^2 \text{Log} L_i(\theta, \eta_i(\theta))}{\partial \eta_i \partial \eta_i} \Big|_{\eta_i = \hat{\eta}_i(\theta)} \frac{\partial \hat{\eta}_i(\theta)}{\partial \theta} \right] \frac{\partial \hat{\eta}_i(\theta)}{\partial \theta}. \end{aligned} \quad (31)$$

¹⁸It is also possible to correct for autocorrelation, as in the case of time-series models, by using a “sandwich estimator” for the covariance matrix.

Appendix C: Evaluation Methodology

To show the importance of dynamic EWSs, it is necessary to scrutinize their forecasting abilities (especially in out-of-sample exercises). To this end, we use the EWS evaluation toolbox developed by Candelon et al. (2012). The main advantage of this framework is that it can be applied to any EWS outputting crisis probabilities, both in-sample and out-of-sample. To be more precise, in a first step we rely on evaluation criteria and comparison tests so as to identify the outperforming model for each country both for *crisis at time t* and *crisis in the following 6 months* events. In a second step, we gauge the ability of the optimal model (static or dynamic) to discriminate between crisis and calm periods by identifying the optimal cut-off for each country. This optimal cut-off is defined as the point between 0 and 1 that best discriminates between crisis and calm periods.

Accordingly, we consider a set of validation methods that take the cut-off into account in the evaluation of EWSs, *i.e.* the *Area Under the ROC criterion* (hereafter *AUC*) and the *ROC* comparison test. These measures lead to a more refined diagnostic than classic EWS evaluation ones, such as the QPS criterion and its corresponding comparison test (Diebold-Mariano, 1995).

The *AUC* is a credit-scoring criterion that reveals the predictive abilities of an EWS by relying on all the values of the cut-off, *i.e.* the threshold used to compute crisis forecasts $\hat{y}_t(c)$, ($c \in [0, 1]$):

$$AUC = \int_0^1 [Se(c) \times (1 - Sp(c))] d(1 - Sp(c)), \quad (32)$$

where $Se(c)$ represents the *sensitivity*, *i.e.* the proportion of crises correctly identified by the EWS for a given cut-off c and $Sp(c)$ is the *specificity*, *i.e.* the proportion of calm periods correctly identified by the model for a cut-off equal to c .

Furthermore, the null hypothesis of the *ROC* test corresponds to the equality of areas under the *ROC* curves associated with the two models under analysis, *i.e.* $H_0 : AUC_1 = AUC_2$. In other words, neither of the models outperforms the other. The alternative hypothesis is the statistical difference of forecasting abilities (the model with the larger *AUC* is the better). The test statistic is

$$W_{AUC} = \frac{(AUC_1 - AUC_2)^2}{\mathbb{V}(AUC_1 - AUC_2)} \xrightarrow{T \rightarrow \infty} \chi^2(1), \quad (33)$$

where $\mathbb{V}(\cdot)$ is an estimator of the variance of the difference between the two areas under the *ROC* curve.

Next, the optimal cut-off is identified for each country by maximizing the Youden index (J), which is an accuracy measure arbitrating between type I and type II errors (misidentified crises and false alarms):

$$c^* = \arg \max_{c \in [0,1]} J(c), \quad (34)$$

where $J(c) = Se(c) + Sp(c) - 1$. A model's ability to correctly discriminate between crisis and calm periods is then given by *sensitivity* and *specificity*. For more details on this evaluation method, see Candelon et al. (2012).

Figures and Tables

Figure 2: Out-of-sample crisis probabilities: crisis at time t , $C1$

Figure 3: Out-of-sample crisis probabilities $C1$ (continued)

Figure 4: Out-of-sample crisis probabilities: crisis in the following 6 months, C_6

Figure 5: Out-of-sample crisis probabilities C_6 (continued)

Table 1: SBC information criterion (time-series logit models, $C1$)

Country	Model 1	Model 2	Model 3	Model 4
Argentina	134.08	118.24	134.45	123.35
Brazil	136.01	109.15	119.000	113.71
Chile	176.44	176.88	180.07	181.27
Colombia	150.31	130.96	156.01	132.12
Ecuador	116.47	108.17	116.00	113.70
Indonesia	64.139	69.478	69.515	75.088
Malaysia	40.197	46.140	45.982	52.035
Mexico	81.994	78.82	86.027	80.462
Peru	131.92	93.243	104.78	91.69
Philippines	139.91	124.35	144.30	128.47
South Africa	144.35	115.78	134.15	120.96
South Korea	69.923	75.002	74.983	82.894
Thailand	40.200	45.96	45.919	51.664
Turkey	94.307	86.812	100.03	86.637
Uruguay	129.77	81.623	73.312	87.141
Venezuela	121.52	117.28	126.48	123.02

Note: Model 1 is the static model (the benchmark), which considers that crisis occurrence can be explained only by macroeconomic factors. Models 2 to 4 are dynamic, including the lagged binary crisis variable y_{t-1} (Model 2), the lagged index π_{t-1} (Model 3), and both of them (Model 4), apart from leading macroeconomic indicators. Bold values correspond to the best model according to SBC for each country. $C1$ denotes the crisis at time t dependent variable.

Table 2: In-sample model comparison (static vs dynamic logit model)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.835	0.864	3.233*	0.072	0.780	0.878	10.715***	0.001
Brazil	0.765	0.903	5.751**	0.016	0.772	0.755	0.803	0.370
Chile	0.754	0.808	2.972*	0.085	0.764	0.789	1.688	0.194
Colombia	0.803	0.851	2.076	0.150	0.656	0.687	2.929*	0.087
Ecuador	0.901	0.914	2.287	0.130	0.898	0.902	0.212	0.645
Indonesia	0.994	0.995	0.124	0.725	0.945	0.947	0.403	0.525
Malaysia	1.000	1.000	0.000	1.000	0.990	0.996	1.575	0.209
Mexico	0.984	0.989	1.168	0.280	0.921	0.943	2.192	0.139
Peru	0.940	0.931	0.214	0.644	0.980	0.980	0.028	0.867
Philippines	0.900	0.942	3.361*	0.067	0.776	0.796	2.892*	0.089
South Africa	0.919	0.952	3.945**	0.047	0.711	0.709	0.099	0.753
South Korea	0.951	0.945	3.917**	0.048	0.827	0.868	9.514***	0.002
Thailand	1.000	1.000	0.000	1.000	0.993	0.993	0.000	1.000
Turkey	0.951	0.971	4.362**	0.037	0.867	0.93	6.865***	0.009
Uruguay	0.938	0.984	10.939***	0.001	0.907	0.916	1.058	0.304
Venezuela	0.928	0.934	0.392	0.531	0.869	0.874	0.818	0.366

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Table 3: In-sample model comparison (time-series vs panel dynamic logit model, $C1$)

Country	AUC criterion			ROC test - All-countries		ROC test - Regional	
	Time-series (y_{t-1})	All-countries Panel	Regional Panel	statistic	p-value	statistic	p-value
Argentina	0.864	0.898	0.901	1.158	0.282	1.219	0.270
Brazil	0.903	0.829	0.832	1.048	0.306	1.152	0.283
Chile	0.808	0.636	0.638	5.033**	0.025	4.996**	0.025
Colombia	0.851	0.845	0.850	0.015	0.902	0.001	0.980
Ecuador	0.914	0.879	0.879	3.509*	0.061	3.503*	0.061
Indonesia	0.995	0.935	0.956	1.186	0.276	1.053	0.305
Malaysia	1.000	0.985	0.980	1.264	0.261	1.287	0.257
Mexico	0.989	0.943	0.943	1.082	0.298	1.077	0.299
Peru	0.931	0.971	0.975	0.606	0.436	0.608	0.436
Philippines	0.942	0.933	0.929	0.216	0.642	0.792	0.374
South Africa	0.952	0.912	0.928	2.217	0.136	1.799	0.180
South Korea	0.945	0.940	0.957	0.056	0.813	1.078	0.299
Thailand	1.000	0.997	0.990	2.068	0.150	1.487	0.223
Turkey	0.971	0.936	0.948	4.397**	0.036	4.440**	0.035
Uruguay	0.984	0.983	0.983	0.134	0.714	0.119	0.730
Venezuela	0.934	0.926	0.929	1.751	0.186	0.508	0.476

See note to table 10.

Table 4: Out-of-sample model comparison (static vs dynamic time-series models)

Country	Crisis at time t : $C1$				Crisis in the following 6 months: $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.997	0.894	0.449	0.503	0.892	0.888	0.099	0.754
Brazil	0.935	0.942	1.482	0.223	0.901	0.950	6.535**	0.011
Chile	0.808	0.816	0.313	0.576	0.763	0.814	8.790***	0.003
Colombia	0.809	0.834	2.812*	0.094	0.631	0.602	0.906	0.341
Ecuador	0.857	0.841	1.224	0.269	0.941	0.856	1.917	0.166
Indonesia	0.899	0.899	0.001	0.979	0.635	0.825	3.672*	0.055
Malaysia	0.945	0.947	0.199	0.655	0.719	0.860	7.260***	0.007
Philippines	0.878	0.902	2.119	0.145	0.915	0.917	0.103	0.749
South Africa	0.905	0.876	0.763	0.382	0.685	0.685	0.004	0.948
South Korea	0.917	0.776	1.442	0.230	0.879	0.820	0.804	0.370
Thailand	0.885	0.830	7.618***	0.006	0.919	0.917	0.253	0.615
Turkey	0.658	0.715	0.384	0.535	0.900	0.965	4.239**	0.040
Uruguay	0.940	0.940	0.002	0.966	0.921	0.923	0.184	0.668
Venezuela	0.842	0.865	3.661*	0.056	0.881	0.834**	5.420	0.020

See note to table 10.

Table 5: Correct predictions (optimal model for each country)

	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	Cut-off	Calm	Crisis	Model	Cut-off	Calm	Crisis	Model
Argentina	0.07	99.43%	100.00%	Static	0.234	98.82%	81.82%	Static
Brazil	0.047	90.45%	100.00%	Dynamic	0.360	94.80%	42.86%	Dynamic
Chile	0.107	84.21%	44.44%	Dynamic	0.273	81.94%	55.56%	Dynamic
Colombia	0.046	75.90%	78.57%	Dynamic	0.092	53.69%	67.74%	Static
Ecuador	0.375	97.66%	66.67%	Dynamic	0.282	90.63%	70.00%	Dynamic
Indonesia	0.031	94.64%	50.00%	Dynamic	0.368	92.59%	66.67%	Dynamic
Mexico	0.150	100.00%		Dynamic	0.149	100.00%		Dynamic
Peru	0.140	100.00%		Dynamic	0.102	96.67%		Dynamic
Philippines	0.083	75.74%	81.82%	Static	0.189	74.68%	84.62%	Dynamic
South Africa	0.077	92.35%	70.00%	Dynamic	0.222	74.84%	52.00%	Static
South Korea	0.003	46.55%	66.67%	Dynamic	0.158	85.09%	68.42%	Dynamic
Turkey	0.069	75.86%	50.00%	Static	0.346	95.21%	76.92%	Dynamic
Uruguay	0.121	99.42%	77.78%	Static	0.215	99.40%	71.43%	Static
Venezuela	0.011	89.47%	66.67%	Dynamic	0.239	77.91%	82.35%	Static

Note: The table displays the percentage of crisis and calm periods (at time t , $C1$ and in 6 months, $C6$) correctly forecast by the optimal model for each country. A crisis period is said to be forecast by a model if the probability of a crisis goes beyond the optimal cut-off at that precise date. The value of the optimal cut-off is calculated for each country on the basis of the in-sample dataset (January 1986 - December 1996). This can be done only for countries that register at least one period of crisis in this time - interval. The optimal model for each country is indicated in the last column.

Table 6: In-sample model comparison (Zhang 2000 crisis definition)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.930	0.952	2.310	0.129	0.811	0.811	0.000	0.988
Brazil	0.854	0.909	1.072	0.300	0.758	0.759	0.047	0.829
Chile	0.986	0.986	0.027	0.871	0.838	0.849	1.531	0.216
Colombia	0.745	0.888	5.026**	0.025	0.861	0.860	1.038	0.308
Ecuador	0.940	0.959	1.811	0.178	0.980	0.981	3.554*	0.059
Indonesia	0.842	0.924	2.833*	0.092	0.749	0.759	0.169	0.681
Malaysia	0.932	0.970	4.377**	0.036	0.747	0.758	0.466	0.495
Mexico	0.892	0.933	1.305	0.253	0.816	0.856	3.026*	0.082
Peru	0.577	0.596	0.467	0.494	0.736	0.756	0.269	0.604
Philippines	0.805	0.877	2.240	0.135	0.613	0.667	0.774	0.379
South Africa	0.892	0.907	4.006**	0.045	0.772	0.783	1.309	0.253
South Korea	0.986	0.99	1.750	0.186	0.994	0.994	0.612	0.434
Thailand	0.687	0.846	2.334	0.127	0.816	0.823	0.371	0.542
Turkey	0.946	0.941	0.140	0.709	0.770	0.786	0.473	0.492
Uruguay	0.804	0.892	2.882*	0.090	0.751	0.788	1.506	0.220
Venezuela	0.929	0.962	3.891**	0.049	0.921	0.921	0.020	0.888

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Table 7: In-sample model comparison (KLR based crisis definition)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.821	0.924	5.938**	0.015	0.797	0.934	16.21***	0.000
Brazil	0.879	0.914	3.790*	0.052	0.827	0.826	0.006	0.936
Chile	0.816	0.826	0.427	0.513	0.775	0.792	0.919	0.338
Colombia	0.882	0.946	5.879**	0.015	0.773	0.797	2.382	0.123
Ecuador	0.948	0.954	2.041	0.153	0.936	0.937	0.072	0.788
Indonesia	0.985	0.985	0.006	0.940	0.938	0.934	0.053	0.818
Malaysia	0.957	0.990	9.281***	0.002	0.847	0.839	0.360	0.548
Mexico	0.957	0.968	0.889	0.346	0.772	0.808	2.395	0.122
Peru	0.944	0.964	1.072	0.301	0.976	0.967	0.566	0.452
Philippines	0.896	0.949	3.467*	0.063	0.752	0.751	0.666	0.415
South Africa	0.913	0.928	2.908*	0.088	0.715	0.754	5.085**	0.024
South Korea	0.985	0.985	0.020	0.886	0.870	0.899	2.683	0.101
Thailand	0.945	0.959	2.004	0.157	0.930	0.94	1.182	0.277
Turkey	0.957	0.962	0.495	0.482	0.912	0.927	1.896	0.168
Uruguay	0.998	0.999	0.667	0.414	0.992	0.996	1.969	0.161
Venezuela	1.00	1.00	0.000	1.000	0.993	0.993	0.000	1.000

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Table 8: In-sample model comparison (Exchange-rate based crisis definition)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.979	0.98	0.002	0.965	0.994	0.995	1.077	0.299
Brazil	0.784	0.962	32.06***	0.000	0.733	0.823	11.25***	0.001
Chile	0.960	0.972	0.919	0.338	0.982	0.987	0.964	0.326
Colombia	0.794	0.940	6.164**	0.013	0.660	0.672	0.183	0.669
Ecuador	0.825	0.940	32.28***	0.000	0.842	0.888	5.113**	0.024
Indonesia	0.960	0.995	11.36***	0.001	0.950	0.987	5.635**	0.018
Malaysia	0.992	0.996	0.288	0.592	0.992	0.991	0.146	0.702
Mexico	0.778	0.965	9.859***	0.002	0.938	0.951	3.690*	0.055
Peru	0.894	0.967	2.770*	0.096*	0.882	0.958	6.125**	0.013
Philippines	0.908	0.992	17.37***	0.000	0.841	0.930	10.41***	0.001
South Africa	0.696	0.904	13.03***	0.000	0.747	0.789	2.509	0.113
South Korea	0.823	0.962	27.89***	0.000	0.772	0.818	6.501**	0.011
Thailand	0.888	0.981	26.01***	0.000	0.941	0.95	2.481	0.115
Turkey	0.769	0.954	49.63***	0.000	0.740	0.818	9.325***	0.002
Uruguay	0.926	0.989	9.421***	0.002	0.880	0.907	2.066	0.151
Venezuela	0.960	0.986	4.815**	0.028	0.902	0.955	11.50***	0.001

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Table 9: Out-of-sample model comparison (Exchange-rates based crisis definition)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.846	0.599	5.107	0.024	0.909	0.922	0.052	0.820
Brazil	0.584	0.901	39.95***	0.000	0.659	0.735	3.871**	0.049
Chile	0.486	0.483	0.006	0.941	0.876	0.877	0.066	0.797
Colombia	0.813	0.784	0.291	0.590	0.804	0.729	1.178	0.278
Ecuador	0.931	0.975	9.887***	0.002	0.896	0.974	15.77***	0.000
Indonesia	0.938	0.975	2.107	0.147	0.932	0.984	6.139**	0.013
Malaysia	0.986	0.983	0.187	0.665	0.993	0.997	1.298	0.255
Mexico	0.509	0.452	0.084	0.772	0.739	0.778	0.214	0.643
Philippines	0.858	0.991	26.32***	0.000	0.852	0.997	20.65***	0.000
South Africa	0.654	0.67	0.033	0.855	0.637	0.684	7.646***	0.006
South Korea	0.785	0.931	7.339***	0.007	0.795	0.793	0.017	0.897
Thailand	0.928	0.953	0.944	0.331	0.929	0.931	0.485	0.486
Turkey	0.750	0.902	11.95***	0.001	0.719	0.84	17.27***	0.000
Uruguay	0.817	0.911	2.127	0.145	0.808	0.87	4.434**	0.035
Venezuela	0.994	0.997	3.758*	0.053	0.989	0.991	0.549	0.459

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Table 10: In-sample model comparison (GDP-based crisis definition)

Country	Crisis at time t , $C1$				Crisis in the following 6 months, $C6$			
	AUC criterion		ROC test		AUC criterion		ROC test	
	Static	Dynamic (y_{t-1})	statistic	p-value	Static	Dynamic (y_{t-1})	statistic	p-value
Argentina	0.701	0.946	28.19***	0.000	0.774	0.872	15.71***	0.000
Brazil	0.851	0.967	25.50***	0.000	0.789	0.848	8.992***	0.003
Chile	0.760	0.991	57.55***	0.000	0.754	0.911	29.05***	0.000
Colombia	0.787	0.965	15.56***	0.000	0.733	0.881	10.74***	0.001
Ecuador	0.651	0.988	80.53***	0.000	0.703	0.910	35.72***	0.000
Indonesia	0.832	0.952	12.26***	0.000	0.828	0.827	0.002	0.968
Malaysia	0.971	0.971	0.000	1.000	0.973	0.979	1.696	0.193
Mexico	0.838	0.990	30.54***	0.000	0.857	0.925	10.61***	0.001
Peru	0.919	0.978	4.674**	0.031	0.877	0.927	5.898**	0.015
Philippines	0.780	0.961	35.47***	0.000	0.878	0.892	1.258	0.262
South Africa	0.745	0.924	24.73***	0.000	0.837	0.839	0.044	0.834
South Korea	0.832	0.945	30.79***	0.000	0.718	0.838	22.65***	0.000
Thailand	0.812	0.986	27.72***	0.000	0.850	0.941	13.91***	0.000
Turkey	0.775	0.945	43.65***	0.000	0.760	0.822	7.378***	0.007
Uruguay	0.902	0.987	16.72***	0.000	0.900	0.929	6.805***	0.009
Venezuela	0.816	0.956	6.424**	0.011	0.756	0.806	4.194**	0.041

Note: The AUC criterion takes values between 0.5 and 1, 1 being the perfect model. Bold values indicate the best model for each country according to the AUC criterion. The null hypothesis of the ROC test is the equality of predictive performance of the two models. Its alternative hypothesis is the statistical difference of the AUC associated with the two models. The asterisks ***, **, and * denote significance at the 1%, 5% and 10% significance levels, respectively. $C1$ ($C6$) denotes the crisis at time t (crisis in the following 6 months) dependent variable.

Figure 6: In-sample crisis probabilities $C6$ Exchange rates-based definition

Figure 7: In-sample crisis probabilities $C6$ Exchange rates-based definition (continued)

Figure 8: Out-of-sample crisis probabilities $C6$ Exchange rates-based definition

Figure 9: Out-of-sample crisis probabilities $C6$ Exchange rates-based definition (continued)

Figure 10: In-sample crisis probabilities $C6$ GDP-based definition

Figure 11: In-sample crisis probabilities C_6 GDP-based definition (continued)