

WP/03/32

IMF Working Paper

Early-Warning Systems: A Survey and a Regime-Switching Approach

Abdul Abiad

IMF Working Paper

Research Department

Early-Warning Systems: A Survey and a Regime-Switching Approach

Prepared by Abdul Abiad¹

Authorized for distribution by Ashoka Mody

February 2003

Abstract

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Previous early-warning systems (EWSs) for currency crises have relied on models that require a priori dating of crises. This paper proposes an alternative EWS, based on a Markov-switching model, which identifies and characterizes crisis periods endogenously; this also allows the model to utilize information contained in exchange rate dynamics. The model is estimated using data for the period 1972–99 for the Asian crisis countries, taking a country-by-country approach. The model outperforms standard EWSs, both in signaling crises and reducing false alarms. Two lessons emerge. First, accounting for the dynamics of exchange rates is important. Second, different indicators matter for different countries, suggesting that the assumption of parameter constancy underlying panel estimates of EWSs may contribute to poor performance.

JEL Classification Numbers: C22, C53, F47

Keywords: Currency crisis; early warning system; regime switching; Markov switching

Author's E-Mail Address: aabiad@imf.org

¹ Economist, Research Department, International Monetary Fund. I am grateful to Roberto Mariano, Lawrence Klein, Tayyeb Shabbir, Ashoka Mody, Andrew Berg, and Sean Campbell for useful comments. An earlier version of this paper was part of my Ph.D. dissertation at the University of Pennsylvania.

| Contents | Page |
|---|------|
| I Introduction..... | 3 |
| II. A Survey of the Post-1997 Empirical Literature on Currency Crises | 7 |
| A. An Assortment of Crisis Definitions..... | 7 |
| B. Differences in Coverage | 9 |
| C. Standard Methodologies, with a Twist | 9 |
| D. New Methodologies | 13 |
| III. A Markov-Switching Approach to Early-Warning Systems | 19 |
| A. Model Specification and Estimation..... | 20 |
| IV. Data Description and Transformation..... | 23 |
| V. Estimation Results | 28 |
| A. Indonesia | 30 |
| B. Republic of Korea..... | 30 |
| C. Malaysia..... | 34 |
| D. Philippines..... | 36 |
| E. Thailand | 38 |
| VI. Forecast Assessment | 38 |
| VII. Conclusions..... | 44 |
| Tables | |
| 1. Early-Warning Indicators..... | 24 |
| 2. Coefficient Estimates for the 22 EWIs (Bivariate Regressions)..... | 26 |
| 3. Final Model Estimates | 29 |
| 4. Speculative Pressure Episodes and Alarm Signals in Indonesia | 31 |
| 5. Speculative Pressure Episodes and Alarm Signals in the Republic of Korea..... | 31 |
| 6. Speculative Pressure Episodes and Alarm Signals in Malaysia | 34 |
| 7. Speculative Pressure Episodes and Alarm Signals in the Philippines | 36 |
| 8. Speculative Pressure Episodes and Alarm Signals in Thailand..... | 38 |
| 9. Forecast Assessment | 41 |
| 10. Measures of Predictive Power | 42 |
| Figures | |
| 1. Indonesia: Crisis Dates, Forecast Probabilities, and Alarm Signals | 32 |
| 2. Republic of Korea: Crisis Dates, Forecast Probabilities, and Alarm Signals | 33 |
| 3. Malaysia: Crisis Dates, Forecast Probabilities, and Alarm Signals..... | 35 |
| 4. Philippines: Crisis Dates, Forecast Probabilities, and Alarm Signals | 37 |
| 5. Thailand: Crisis Dates, Forecast Probabilities, and Alarm Signals | 39 |
| Appendix I: A Summary of the Recent Empirical Literature on Currency Crises | 46 |
| References..... | 54 |

I. INTRODUCTION

A succession of currency crises in the past decade has led to a proliferation of theoretical and empirical papers on the factors that brought about these crises. Several papers have also focused on the issue of anticipation—devising early warning systems that give policymakers and market participants warning that a crisis is likely to occur. Two approaches to constructing early warning systems have become standard: limited dependent variable probit/logit models and the indicators approach of Kaminsky, Lizondo, and Reinhart (1998, hereinafter referred to as KLR). Berg and others (2002) assess the performance of these models and find that they have outperformed alternative measures of vulnerability such as bond spreads and credit ratings. However, while these models are able to anticipate some crises, they also generate many false alarms.

There are several well-known methodological issues associated with the existing early warning models. Perhaps the most significant is that they require an a priori dating of crisis episodes before they can be estimated. The most common procedure for doing so is by taking changes in exchange rates, reserves, and/or interest rates; choosing weights for each; and combining them into an index of speculative pressure, specifying a sample-dependent threshold, and identifying crises based on whether or not the index exceeds the threshold. But as is evident from the survey of 26 recent empirical studies of currency crises in Section II below, this simple procedure has been applied in a multitude of ways, resulting in different periods being identified as crises.²

The threshold procedure provides a set of crisis dates, but raises even more problems. First, the choice of the crisis-identification threshold is arbitrary. A selected sampling of thresholds used in the literature include the threshold of $1.5 \times \sigma$ (where σ is the sample standard deviation) used in Aziz and others (1999), $1.645 \times \sigma$ in Caramazza and others (2000), $1.75 \times \sigma$ in Kamin and others (2001), $2.5 \times \sigma$ in Edison (2000), and $3 \times \sigma$ in KLR. Different choices of threshold will obviously result in different crisis dates and different estimated coefficients. Moreover, the threshold is sometimes treated as a free parameter and chosen so that the fit of the model is maximized (Kamin and others, 2001), or so that a set percentage, say 5 percent, of all observations are crises (Caramazza and others, 2000).

Second, the sample-dependent nature of the threshold definition implies that future data can affect the identification of past crises. Thus one can observe cases of disappearing crises, as documented by Edison (2000). Since the threshold is defined in terms of the sample standard deviation, the occurrence of a new, relatively large crisis, such as the Asian crisis, results in

² For example, Kamin, Schindler, and Samuel (2001) compare their identified crisis dates with those identified by KLR and find that only 61 percent of crisis dates were commonly identified.

previously identified crises no longer being identified as such. Edison notes that the threshold methodology identifies five crises in Malaysia using pre-1997 data, but these all disappear, and only one crisis is identified (the 1997 crisis itself), when data up to 1999 are included in the sample.

Third, many of these studies make ad hoc adjustments to the binary crisis variable that may introduce artificial serial correlation. One common procedure is the use of “exclusion windows,” which omits any crises identified by the threshold method if they follow a previous crisis within a certain window of time. As is the case with the threshold level, the width of the exclusion window is arbitrary, and has been chosen to be anywhere from one quarter (Eichengreen, Rose, and Wyplosz, 1996) to as long as 18 months (Aziz and others, 1999) and even three years (Frankel and Rose, 1996, using annual data). The motivation for using exclusion windows is to eliminate identifying speculative pressure episodes as new crises if they are just continuations of previous ones. But in doing so, one eliminates any information the sample contains regarding crisis duration. More seriously, it introduces artificial serial correlation in the dependent variable that few studies account for. The estimated probit/logit models implicitly assume independence across observations t . But using an exclusion window means that $C_t = 1 \Rightarrow \Pr(C_{t+j} = 1) = 0$ for $j=1,2,\dots,J$, where J is the width of the exclusion window.³

Finally, information is lost when transforming a continuous variable into a binary variable. In particular, potentially useful information on the dynamics of the dependent variable is discarded. The critique regarding information loss can also be made regarding the treatment of the indicators in the KLR approach, where the explanatory variables themselves are transformed into binary signals.

Given these problems, is there an alternative approach? This paper proposes an EWS methodology, based on a Markov-switching model with time-varying transition probabilities, that can address these issues. First, the model does not require a priori dating of crisis episodes; instead, identification and characterization of crisis periods are part of the model’s output, estimated simultaneously with the crisis forecast probabilities in a maximum-likelihood framework. One thus avoids the pitfalls associated with the threshold dating procedure described above. Additionally, by exploiting information in the dynamics of the dependent variable itself, the model is better able to send warning that a significant exchange rate adjustment is likely.

³ Another procedure that introduces artificial serial correlation, used by KLR (1997) and Berg and Pattillo (1999a) among others, is the practice of setting the dependent variable equal to one in the 24 months *preceding* crises identified using the threshold method. The rationale behind the procedure is to improve model fit for variables that exhibit abnormal behavior in the periods leading up to a crisis.

The assumptions that underlie a Markov-switching model are both concise and intuitive. The first assumption is that there are two states: tranquil periods and speculative attack periods. But we do not directly observe these states; that is, this binary “crisis” variable is *latent*. This brings us to our second assumption: there are directly observable variables whose behavior changes depending on the value of the crisis variable. Most obviously, the behavior of exchange rates is different during periods of speculative pressure than during tranquil periods.⁴ In particular, we expect much greater exchange rate volatility as well as higher average depreciations during speculative attacks. Finally, we assume that given the current state—tranquil or crisis—there is a certain probability of staying in the same state, or of moving to the other state. In our model, the probability of moving from the tranquil state to the crisis state depends on the strength or weakness of a country’s fundamentals.

Several studies have used Markov-switching models in developing theoretical models of speculative attacks. Jeanne and Masson (1998) and Fratzscher (1999) develop currency crisis models with multiple equilibria and use a Markov-switching variable to model switches between these equilibria. In both cases, however, the probability of switching from one equilibrium to another is constant. In contrast, the model in this paper allows switches from the optimistic, no-attack equilibrium to the pessimistic, speculative attack equilibrium to be a function of various indicators.

Two other papers have used Markov-switching with time-varying probabilities to empirically model currency crises. Cerra and Saxena (2002) use a Markov-switching model to look at the 1997 Indonesian crisis and investigate whether the crisis was due to domestic factors, monsoonal factors, or pure contagion from neighboring countries. Their model differs from the one explored here, mainly because the only variable that affects the time-varying probability in their model is a measure of contagion, based on exchange market pressure in neighboring countries. Fundamentals in their model affect only the mean of the exchange rate. In contrast, our view is that domestic and external fundamentals affect the probability of a crisis occurring and, hence, should enter into the time-varying probability equation rather than affecting only the level of the exchange rate.

The most closely related work is by Martinez-Peria (2002), who also estimates a Markov-switching model with time-varying probabilities to model speculative attacks on the European Monetary System (EMS), using data from 1979 to 1993. That paper evaluated the ability of the Markov-switching model to identify crisis episodes,⁵ and assessed the degree to

⁴ One can substitute the speculative pressure index for the exchange rate if unsuccessful speculative attacks are also of interest; see Section IV below.

⁵ In this regard, the results are positive: Martinez-Peria finds that the Markov-switching model is able to identify all the crisis episodes identified by the methods used by Eichengreen, Rose, and Wyplosz (1996) but also identifies 25 additional crisis episodes. She

which five variables—domestic credit growth, the import-export ratio, the unemployment rate, the fiscal deficit, and interest rates—determined crisis vulnerability in the EMS. We extend this work and focus primarily on the use of the model as an early-warning system. First of all, we begin by looking at a wider set of 22 early-warning indicators. In addition to the standard macroeconomic indicators used in other early-warning systems, we also explore indicators relating to the characteristics of capital flows and to financial sector soundness. Second, the predictive ability of the model is assessed both in sample and out-of-sample. Finally, the Martinez-Peria study assumed that the parameters of the model were uniform across countries and pooled the data to get parameter estimates. This is probably an innocuous assumption for the set of advanced economies in her study, which are broadly similar. But for developing countries, such an assumption might not hold. If a country is relatively more open or has fewer capital controls than other countries, for example, the coefficients on measures of external imbalance may be larger. In this paper, the model is estimated separately for each of the five Asian crisis countries (Indonesia, Korea, Malaysia, the Philippines, and Thailand).

With the usual caveat that all early-warning systems are far from perfect and serve only to synthesize information and supplement more in-depth country knowledge, the model does a good job of anticipating crises. It correctly anticipates two-thirds of crisis periods in the sample and, just as important, sends many fewer false alarms than existing models. In the January 2000–July 2001 out-of-sample period, no warning signals are sent for three of the countries (Korea, Thailand, and Indonesia), but vulnerabilities were signaled for Malaysia and the Philippines in mid-2001, mainly owing to a decline in competitiveness and a slowdown in exports.

The other significant contribution of this paper is a detailed survey of the recent empirical literature on currency crises. Covering 30 studies written since 1998, the survey is meant to complement the survey of the pre-1997 empirical literature provided by KLR. We contrast the studies in terms of the crisis definition they use; the geographical coverage, time period, and frequency of their datasets; and the methodologies used. This survey can be found in Section II. Section III describes the Markov-switching model with time-varying probabilities in detail. The data used in the estimation are described in Section IV, and Section V presents the estimation results and a country-by-country analysis. Section VI assesses the model's predictive ability both in sample and out-of-sample, and Section VII concludes.

then finds evidence in news reports and central bank releases that 21 of these 25 periods were indeed speculative attacks.

II. A SURVEY OF THE POST-1997 EMPIRICAL LITERATURE ON CURRENCY CRISES

This section summarizes the results of 30 selected empirical studies of currency crises written since 1998.⁶ The table in Appendix I provides a summary of the studies. The first row lists the datasets used in the studies, in terms of country coverage, frequency of the data used, and the sample period. The second row gives the definition of currency crisis used in the study. The remaining rows describe the indicators examined, the methodological approach used, and the main findings of each of the studies. The survey is organized thematically, first comparing the various crisis definitions used, followed by a discussion of their datasets, and concluding with a more thorough discussion of methodology, indicators used, and the main conclusions.

A. An Assortment of Crisis Definitions

KLR already noted a large variation in the way “crisis” is defined in the pre-1997 studies. If anything, this variation has increased in recent studies, as can be seen in the second row of the Appendix table. Only nine of the studies surveyed use the speculative pressure index in what has come to be considered the standard form, consisting of a weighted average of nominal exchange rate, reserve and interest rate changes (the latter frequently dropped due to lack of market-determined data), and converted into a binary crisis variable using a threshold specified using a sample-dependent standard deviation.⁷ Even among these, there is considerable variation with regard to the inclusion of interest rate changes, the weighting of the various components, the threshold used to define the binary variable, and the treatment of high-inflation episodes.

Four other studies use variants of the speculative pressure index approach. Burkart and Coudert (2000) combine the KLR crisis dates with those identified using the dates of Milesi-Ferretti and Razin (1998), which are based solely on large and accelerating exchange rate depreciations, and further amend these dates using “expert judgment.” Ghosh and Ghosh (2002) concentrate solely on “deep” currency crises, which they define as crises resulting in a decline in the GDP growth rate of at least 3 percentage points. Kamin, Schindler, and Samuel

⁶ Private sector models, such as Goldman Sachs’ GS-WATCH, the Deutsche Bank Alarm Clock (DBAC), and the Emerging Markets Risk Indicator of CSFB are not covered in this survey, but a brief description of these models can be found in Berg and others (2002). We also omit the KLR paper itself and its elaborations, which are documented in a book-length treatment in Goldstein, Kaminsky, and Reinhart (2000).

⁷ These include Aziz, Caramazza, and Salgado (1999), Berg and Pattillo (1999), Caramazza, Ricci, and Salgado (2000), Collins (2001), Edison (2000), Herrera and Garcia (1999), Krkoska (2001), Nag and Mitra (1999), and Weller (1999).

(2001) use *real* as opposed to nominal exchange rate changes in the speculative pressure index, eliminating the need to treat high-inflation countries separately. And Zhang (2001) avoids averaging and weighting issues altogether, by treating exchange rate and reserve changes separately, identifying crises when either of the two crosses a sample-dependent threshold.

Seven of the studies focus exclusively on successful speculative attacks, which result in a significant depreciation.⁸ But there is even greater variation within this class, as regards whether to use real or nominal depreciations, and how large and rapid the depreciation must be to qualify as a crisis.

Finally, nine of the studies eschew the use of a binary crisis variable altogether.⁹ The primary rationale for doing so is the loss of information that comes from transforming a continuous variable into a binary one. Why should speculative pressure just below the threshold be coded the same as a completely tranquil period?¹⁰ And why shouldn't one differentiate between moderate speculative attacks which are just above the threshold and extreme ones which far exceed it? Thus, four of these studies use the continuous speculative pressure index as their dependent variable. The remaining two transform the index into a score with a bounded range: in Hawkins and Klau (2000), the crisis measure is discrete and ranges from -10 to 10, while in Eliasson and Kreuter (2001) an extreme-value distribution is fitted to tail-end observations and its cumulative distribution function (cdf), which is bounded between 0 and 1, is used as the continuous dependent variable.¹¹

⁸ These include Apoteker and Barthelemy (2001), Bruggemann and Linne (2000), Bussiere and Mulder (1999a, b), Esquivel and Larrain (1998), Grier and Grier (2001), Kaufmann, Mehrez, and Schmukler (2000), Kumar, Moorthy, and Perraudin (2002), and Osband and Van Rijckeghem (2000). Mulder, Perrelli, and Rocha (2002) use both the continuous crisis index and the standard threshold variable.

⁹ These include Cerra and Saxena (2002), Eliasson and Kreuter (2001), Hawkins and Klau (2000), Kwack (2000), Nitithanprapas (2000), and Vlaar (2000).

¹⁰ Flood and Marion (1999) note that to the extent that a speculative attack is anticipated, the size of jumps in exchange rates, interest rates, and reserves at the attack point is lessened and "selecting only extreme values of the speculative pressure index...may reduce the share of predictable crises in the sample."

¹¹ See below for a more detailed description of their crisis intensity index.

B. Differences in Coverage

The empirical studies surveyed differ widely in regard to their geographical coverage. All of the studies except one use a panel of countries, and all focus on emerging market economies, with only a fourth of the studies also including advanced economies in their datasets. And although most of the studies have a global orientation, trying to draw lessons that can be generalized across regions, seven of the studies focused on specific regions—Bruggemann and Linne (2000) and Krkoska (2000) study transition economies, Herrera and Garcia (1999) look exclusively at Latin American economies, and Cerra and Saxena (2002), Kwack (1999), Nag and Mitra (1999), and Zhang (2001) analyze the Asian crisis countries.

Unlike some of the studies surveyed in KLR which investigate crises as far back as the early 1950s, all of the recent papers focus only on the last three decades, and in fact only one-fourth have data going back as early as the 1970s. The remainder are split evenly between those that use data from the 1980s onwards, and those that focus exclusively on the crises in the 1990s. Data availability was the primary determinant of the time periods covered in the studies, and also determined the frequency of the data. Unsurprisingly, most studies indicate a clear preference for higher frequency data, both for a more precise dating of crisis episodes and also since the usefulness of an early-warning system depends largely on the timeliness of the data it is based on. Sixteen of the studies use monthly data, and another three use quarterly data. The seven studies that rely on annual data do so either because their focus is on a cross-sectional comparison, or because they utilize special data which is unavailable at a higher frequency. Two papers that fall into the latter category are Ghosh and Ghosh (2002), who look at the impact of institutional variables on crisis likelihood, and Kaufmann, Mehrez, and Schmukler (2000), who use an annual survey of local managers to gauge the extent to which they have private information about a country's vulnerability to crisis.

We now turn to a more detailed description of the various studies, which can be grouped into two categories based on the methodology employed. The first group consists of studies which use one of the two standard approaches in the literature—either the indicators approach or limited dependent variable probit/logit modeling—but introduce some novelty, such as a new set of indicators or a new method for assessing performance. The other group consists of studies which utilize methodological alternatives to the standard approaches.

C. Standard Methodologies, with a Twist

Roughly half of the studies use one of the two standard methodologies, but introduce a novel twist in the form of a new set of indicators, an alternative transformation for the indicators, a

different crisis definition, or a new approach to assessing the results.¹² Berg and Pattillo (1999) revisit the indicators model of KLR, adding two new indicators—the ratio of M2 to reserves, and the current account to GDP ratio—and updating the data through 1996 to see whether the Asian crisis could have been anticipated. They then compare this predictive performance with three probit-based alternatives: one that uses the KLR signals as right-hand side inputs, a second that uses piecewise linear functions (which allows for a more general nonlinearity than the step function used in KLR), and a third using a simple linear specification for the variables. They find that the probit model which uses untransformed variables does better than both the signals-based and piecewise linear probit models in predicting the Asian crisis out-of-sample. The piecewise linear probit model fits better than the KLR probit model in sample, but its poorer out-of-sample performance indicates that this is probably due to overfitting. Berg and Pattillo are also very thorough in comparing the models, using various measures of predictive power including various accuracy scores such as the quadratic probability score, as well as goodness-of-fit based on the percentage of correctly called crises relative to false alarms.

Edison (2000) also provides a very thorough assessment of the KLR methodology, by updating and conducting out-of-sample forecasts, and by submitting it to a battery of various sensitivity tests. She finds that the performance of the various indicators is reasonably robust, although there are differences in the performance of various indicators across regions. She also identifies some shortcomings of the various crisis identification methods, including a documentation of disappearing crises due to the use of sample-dependent thresholds. Finally, she examines the usefulness of the model for monitoring a single country over time, using Mexico as an example, and for doing cross-country vulnerability comparison at a point in time.

A third paper which, along with Berg and Pattillo (1999) and Edison (2000), sets the standard for the proper assessment of early-warning systems, is Kumar, Moorthy, and Perraudin (2002), who use a standard logit model to construct a predictive model for currency crashes. They focus solely on successful speculative attacks, and hence use only exchange rate changes to define crisis episodes. They use various accuracy scores, goodness-of-fit tables, and a much larger hold-out sample (one-third of total observations) for out-of-sample evaluation. One of the non-statistical methods they employ for evaluating their model is whether a trading strategy based on its predictions can be profitable.

¹² Two of the studies included in Table A.1, Aziz, Caramazza, and Salgado (1999) and Esquivel and Larrain (1998), have a methodology and results that are very similar to other studies so we do not discuss them in detail here.

Kamin, Schindler, and Samuel (2001) use annual data on 26 emerging markets from 1981–1999 to investigate the relative influence of domestic and external factors in bringing about currency crises. In contrast to other studies, they construct their speculative pressure index using real instead of nominal exchange rate changes, eliminating the need for treating high-inflation episodes separately. Five domestic indicators, four external balance indicators, and three external shock indicators are used. Although they use a standard probit approach with indicators that have been explored in other studies, their contribution lies in their decomposition of probability changes into those brought about by domestic factors and those due to external factors. They find that external balance and external shock variables contribute little to the average probability of crisis, but account for a significant portion of the increases or spikes in the probabilities during the crisis years themselves. They interpret this as evidence that domestic factors are still the primary determinants of vulnerability to crisis, but that it is external factors that often push vulnerable countries over the edge.

The next group of studies explore new measures and explanatory variables that contribute to vulnerability. Caramazza, Ricci, and Salgado (2000) investigate the importance of trade and financial linkages in the transmission of currency crises across countries. Critiquing the trade linkage measures used by Glick and Rose (1998), they construct a new measure of trade linkages based on both the price effects (via the REER) and income effects (via growth slowdowns in partner countries) due to crises in other countries.¹³ Financial linkages are measured both through common creditor channels (using data from BIS-reporting banks), as well as by stock market correlations with the original crisis country. After controlling for standard economic factors such as overvaluation, current account imbalances and output growth, they find that financial linkages in the form of a common creditor channel significantly increase the likelihood of contagion. Trade linkages, on the other hand, are found to be (marginally) insignificant. However, when the trade linkage measures are interacted with the current account, the interaction term becomes significant, indicating a strong trade spillover effect for countries with already weak external positions.

Bussiere and Mulder (1999a) study the effect of political instability on crisis depth, the latter being measured by a weighted average of exchange rate and reserve changes. To do this, they supplement Tornell's (1999) cross-section analysis of the Tequila and Asian crises with four measures of political instability. These include two measures of fragmentation (in the legislature and in the ruling coalition), one measure of electoral uncertainty, and dummy variables for election dates. They find the first two to be insignificant and the latter two to be significant and robust. Interestingly, they find that post-election periods, even more than pre-election periods, bring about greater vulnerability. A second paper (Bussiere and

¹³ See Appendix I.C, pp. 43–44 of Caramazza and others (2000) for more details.

Mulder 1999b) finds that the presence of an IMF-supported program significantly reduces crisis depth, even beyond any positive effect a program might have on included fundamentals.

A third, more recent paper (Mulder, Perrelli, and Rocha 2002) investigates the role of corporate balance sheet indicators, as well as governance standards as measured by the strength of creditor and shareholder rights. These indicators are tested in both the probit model of Berg and Pattillo (1999) to assess their impact on crisis likelihood, as well as in the Sachs, Tornell, and Velasco (1996) cross-sectional analysis to assess their impact on crisis depth. They find that high leverage and short maturity structures increase both the likelihood of crises as well as the depth of crises. The impact of these balance sheet variables is greater when bank credit to the corporate sector is large, suggesting that corporate weaknesses are transmitted through the banking sector. Shareholder rights also have a large impact on crisis probabilities.

Another study that looks at new explanatory variables is Kaufmann, Mehrez, and Schmukler (2000), who explore whether local managers have an informational advantage over other market participants. They first look at descriptive evidence from financial markets. Mutual fund holdings did not decrease, except in Malaysia, before the Asian crisis, indicating that the crisis was not anticipated by these funds. Similarly, BIS bank lending did not decrease, although BIS bank deposits by residents of Korea and Thailand increased pre-crisis, indicating capital flight in anticipation of the crisis. There was some evidence that currency forecasters expected a slight depreciation in Thailand, but not in the other Asian crisis countries. Finally, credit rating downgrades followed rather than preceded the crisis. They then examine more direct evidence of local managers' views, taken from the Global Competitiveness Survey. Here as well, there is evidence of deteriorating local sentiment in Thailand in December 1996, but not in other countries. Kaufmann, Mehrez, and Schmukler undertake an econometric analysis by extracting the private information of local managers—measured by the residual from an ordered probit regression of survey opinions on macroeconomic variables—and find that this private information is a significant predictor of exchange rate volatility.

Weller (2001) examines whether financial liberalization affects the degree to which a country is vulnerable to a currency crisis. He does this first by comparing tranquil and crisis period means of various indicators, both before and after financial liberalization, as well as comparing pre- and post-liberalization crisis periods directly. He then runs separate logit regressions on the pre- and post-liberalization subsamples, and finds significant differences between the two estimates. Sensitivity to changes in vulnerability indicators—especially short-term loans to reserves and real overvaluation—increases significantly after liberalization.

Grier and Grier (2001) investigate whether the exchange rate regime of a country affects the degree of depreciation, or the returns of the stock market. Examining a cross-section of

25 countries in 1997, they find that after controlling for other factors, countries with a peg depreciated more than those without one. Countries with a peg also had lower stock returns. Finally, Kwack (2000) investigates how much of a role external factors (as measured by LIBOR) and financial fragilities (as measured by non-performing loan ratios and corporate leverage ratios) played in the Asian crisis, by running a panel regression on annual 1995–1997 data from seven Asian countries. Despite the limited number of observations (fourteen) in the OLS regression, he finds the LIBOR and the nonperforming loan ratio significant in explaining the variance of the exchange market pressure index. A second regression finds that the non-performing loan ratio can be explained by the debt-equity ratio in the corporate sector. These two results are then combined into a semi-reduced form OLS regression, where the crisis index is significantly related to LIBOR and to the corporate debt-equity ratio; these two alone are found to explain 85 percent of the variance in the crisis index.

Bruggemann and Linne (2000) investigate whether the indicators approach can be used to cover a qualitatively very different set of countries—the transition economies, more specifically the EU accession countries plus Russia and Turkey. In addition to assessing the usefulness of conventional indicators for these economies, they also analyze indicators of capital flight risk and banking sector fragility. They find that the standard macroeconomic variables, particularly dwindling reserves, an overvalued exchange rate and a rising budget deficit, are useful predictors for the transition economies as well, but they also find that some banking sector indicators—the ratio of lending to deposit rates, and the size of bank deposits relative to GDP—also have predictive power.

The final study in this group, Eliasson and Kreuter (2001), uses a standard limited dependent variable (logit) methodology, but they question the use of a binary crisis measure and suggest an alternative continuous crisis measure based on a five-step procedure. This entails fitting an extreme-value distribution to the tail end of exchange rate changes, and using the c.d.f. of the distribution as a bounded measure of crisis intensity. The same procedure is applied to interest rate changes and to deviations of the interest rate from its long-term mean, and the maximum of the three c.d.f. values is used, thus avoiding aggregation via a weighted average. Finally, Eliasson and Kreuter smooth the crisis intensity index via an exponentially weighted moving average over the last six observations, with the largest of the six observations getting the residual weight. The crisis measure obtained is used as the dependent variable in a multinomial logit model. They find that the continuous crisis variables provide a more accurate description of the crises in the 1990s than the binary crisis variable, and that their estimated model performs well in explaining these events.

D. New Methodologies

Almost half of the studies in this survey explore alternative econometric specifications to the standard indicators and probit/logit models. Despite the variety of approaches proposed, most of these studies identify the same set of shortcomings in the existing approaches—which we

have already enumerated in the introduction—to justify an exploration of alternatives. But few, if any, of the proposed alternatives are able to address all these shortcomings. Several still rely on the crisis dating methodology of the standard approaches. Others are more successful at addressing the weaknesses of the standard models, but have problems of their own. What should be clear after surveying these new approaches is that although none of the models is perfect, each has its own strengths, and an awareness of the advantages and disadvantages of each model is essential so that practitioners can select the proper model for a given task. Indeed, one of the objectives in compiling this survey was to increase awareness of the existence of these models, many of which have remained as unpublished working papers. We now turn to a discussion of the individual studies.¹⁴

Nag and Mitra (1999) use an **artificial neural network** (ANN) to construct an early-warning system for currency crises, and compare its performance to the indicators approach using monthly data for Indonesia, Malaysia and Thailand from 1980–1998. The primary advantage of ANNs are their flexible specification, and their ability to capture complex interactions among variables. However, this flexibility can also be a potential drawback, as the danger of overfitting is much greater than in other methodologies, given the large number of variables and the ability of the neuron layers to fit the data. Another drawback is the “black box” nature of ANNs. Because there are no coefficient estimates, and the interactions among the variables can be very complicated, it is difficult to determine which indicators are behaving abnormally and driving the forecast probabilities.¹⁵

In their country-by-country replication of the KLR approach using sixteen indicators, Nag and Mitra find that different indicators are useful for different countries; somewhat surprisingly, they rarely find real overvaluation to be a significant indicator. They then estimate a different ANN model for each country, using a genetic algorithm to train the models. Because lags of anywhere from zero to four quarters are also allowed, the final models for each country have a large number of variables, from 13 in Indonesia to 23 in Malaysia. Inadequate information is given regarding the construction of the ANN, such as the number of hidden layers and hidden neurons, the transformation function used, or the parameters of the genetic algorithm used to train the ANN. They do not report the in-sample model forecasts, which is estimated up to end-1996, but their out-of-sample results show very high crisis probabilities—in the vicinity of 80 percent—for all three countries in the

¹⁴ The two empirical studies using a regime-switching approach, Cerra and Saxena (2002) and Martinez-Peria (2002), were discussed in the Section I and hence are omitted from this section.

¹⁵ Appendix I of Nag and Mitra (1999) provides a brief description of neural networks. For a thorough survey of neural networks intended for econometricians and economists, see Kuan and White (1994) and the comments that follow, especially Lewbel (1994).

months before the crisis broke. A more thorough evaluation is needed for this potentially very promising approach.

Collins (2001) uses a **latent variable threshold model** to study the timing of currency crises. She assumes that a crisis occurs when some unobservable process crosses a threshold, where the latent variable is assumed to follow Brownian motion with drift. Conditional on the drift factor, the distance to the threshold and the variance of the Brownian motion, the probability of crisis occurrence has an inverse Gaussian distribution. The distance and drift factors are modeled as linear functions of five standard indicators, with overvaluation, the CA/GDP ratio and short-term debt to reserves affecting distance to the threshold, and export and reserves growth affecting the drift factor. Estimating the model using monthly data for 25 emerging markets from 1985–1988, she finds that short-term debt influences distance, while reserve growth influences the drift factor. Interestingly, the drift factor is always negative, indicating that the process is moving away from the crisis threshold over time. Collins also tests the model against two alternative specifications—a probit model and a Poisson model—and finds that the threshold model is found to fit better than either alternative.

Blejer and Schumacher (1998) propose using a **value-at-risk** approach to analyze central banks' balance sheets and assess solvency risk, which can reduce the credibility of an exchange rate peg. They consider exposure to several risk factors, including volatility in exchange rates, international interest rates, and country risk. The methodology is developed in detail, but Blejer and Schumacher refrain from estimating the model using existing data, instead laying down guidelines on how this approach might be made operational.

Vlaar (2000) also takes a different methodological approach to early-warning systems. First, he uses the continuous crisis index itself instead of a binary crisis dummy. Second, he models the crisis index as being drawn from a **mixture of two normal distributions**. The mean and volatility of each distribution is modeled as a function of various indicators, as is the relative weighting of the two distributions. He finds that inflation, overvaluation and reserve losses are the most significant determinants of vulnerability, but in addition, he finds a substantial degree of dynamics—past exchange rate and reserve deterioration and volatility are themselves significant predictors of future vulnerability. Regional exchange rate volatility also affects vulnerability, suggesting the possibility of contagion. Both in-sample and out-of-sample, the model is able to predict about 7 of 8 crises, but at the expense of sending false alarms through approximately a third of tranquil periods.

Zhang (2001) proposes using the **autoregressive conditional hazard (ACH)** model developed by Hamilton and Jorda (2002). In essence, it is simply an addition of a new explanatory variable—the duration of the last tranquil period—and an alternative functional

form for the crisis probability: $\Pr(Crisis_{i,t} = 1) = \frac{1}{C + \alpha d_{i,t-1} + \gamma' Z_{i,t}}$, where d gives the duration or length of the last tranquil period and Z is a vector of standard indicators.¹⁶ Using monthly data for Indonesia, Korea, the Philippines and Thailand from 1993–1997, Zhang estimates both a probit model and the ACH model and asserts that the ACH model fits the data better, based on a comparison of log likelihoods; however, such a comparison is invalid since the models are not nested. Further, he finds that *none* of the standard vulnerability indicators is significant, but that the duration variable is highly significant. A contagion measure is also constructed, based on the duration of the most recent tranquil period in *any* of the four countries, and not only is it significant, but it also drives out the effect of the domestic duration variable.

It is difficult to assess the generality of Zhang’s findings, since the model is estimated on a very short sample. Not surprisingly, the fitted model is able to predict the crises in Korea, Indonesia and the Philippines on the basis of the onset of the crisis in Thailand. Zhang identifies May 1997 as the starting date of Thailand’s crisis, since he adopts a different crisis dating methodology, treating exchange rate and reserve changes separately and using a three-year moving window for computing the standard deviation threshold.

Adopting a more standard econometric approach, Krkoska (2001) estimates a **restricted VAR** to analyze vulnerability in four transition countries. An index of speculative pressure is constructed, and the continuous index enters the VAR along with four other endogenous variables—the real exchange rate, industrial production, FDI, and the current account. Five exogenous variables also enter the specification—the CA-FDI gap, growth in real domestic credit, inflation, the DM-US\$ exchange rate, and industrial production in the EU. Given the limited data (quarterly data from 1994–1999) and the large number of parameters, many ad hoc restrictions were imposed for the VAR to be tractable. Krkoska finds that the gap between the current account and FDI is the most significant predictor of crisis vulnerability; in all instances that this gap exceeded 5 percent of GDP, a crisis ensued. Overvaluation and a slowdown in the EU are also found to be significant predictors of vulnerability.

Burkart and Coudert (2000) utilize **Fisher discriminant analysis** in their analysis of currency crises. Discriminant analysis aims to classify a dependent variable into one of K given states, based on information from a set of predictor variables. The principle underlying the analysis is to determine whether the K states differ with regard to the mean of a variable,

¹⁶ Following Hamilton and Jorda, Zhang uses a “smoothing transform” function to ensure that the probability lies between 0 and 1.

and then to use that variable to construct predictions for the states.¹⁷ In this study, the sample is divided into $K=2$ states, where the four quarters that precede each crisis are one state (“crisis”), and all other periods are the other state (“tranquil”). Burkart and Coudert start with a set of 34 potential indicators and reduce this set to six indicators—the ratio of reserves to both M2 and debt, the ratio of short-term debt to total debt, real overvaluation, inflation, and a regional contagion indicator—based on performance. The find that the discriminant model based on these indicators yields relatively good performance: four out of five crises are predicted correctly and only one out of five non-crises result in false alarms.

In contrast to the other studies surveyed here, Hawkins and Klau (2000) are not interested in the estimation of a model. Rather, their objective is to **present vulnerabilities in a simple, transparent manner**. They do this by transforming various indicators of external and banking sector vulnerability into discrete scores, where higher scores indicate increased vulnerability. For example, four continuous indicators for external vulnerability are individually transformed into discrete measures, taking on five possible values: -2, -1, 0, 1 and 2. These scores are then summed, using weights of 1.25 so that the score falls within an interpretable range of -10 to 10. A similar transformation is also done to compute speculative pressure itself, using reserve changes, real interest rate changes, and exchange rate changes at both the 3-month and 12-month horizons. But in its essence, the Hawkins and Klau proposal is just a variant of the KLR approach. The only differences are that (i) they classify indicators into five categories, as opposed to the binary 0–1 categories of KLR; (ii) for the speculative pressure index, they categorize and then sum, whereas KLR sum first and then categorize into a 0–1 crisis dummy; and (iii) Hawkins and Klau sum their indicators using equal weights, while KLR weight their indicators based on predictive performance. Finally, although a five-category variable is more informative than the binary transformation done in the signals approach of KLR, the thresholds used to create the five categories are arbitrary, unlike the threshold used in KLR which is chosen to maximize the signal-to-noise ratio.

Several of the studies surveyed emphasize the importance of potential interactions among various indicator variables. Nitithanprapas and Willett (2000) are critical of other empirical studies for failing to account for these interactions. They note, for example, that current accounts which due to overvaluation are worrisome, while those that are due to inflows of FDI are probably less dangerous. Using a **slope dummies regression** approach, they test the Lawson dogma that current account deficits are worrisome only if caused by a fiscal deficit, as well as other hypotheses linking current account deficits to real overvaluation and to FDI

¹⁷ The idea is to find a linear combination of the indicators that exhibits the largest difference in the group means relative to the within-group variance. This linear combination is known as Fisher’s discriminant function, or the first canonical variate. A cut-off point can then be used to generate predictions.

inflows. They also test whether lending booms increase vulnerability to a crisis, and whether adequate reserves (measured against both M2 and short-term debt) mitigate these risks. They find no support for the Lawson dogma at least in the context of their sample (the Mexican and Asian crises), but they do find evidence of interactions between the current account, overvaluation and FDI. They also find the effect of lending booms significant, and that adequate reserves lessen crisis vulnerability.

In addition to addressing the issue of interactions in their study, Ghosh and Ghosh (2002) introduce several additional innovations. They focus exclusively on “deep” currency crises, defined as currency crises—identified using a standard speculative pressure index and threshold—which result in an appreciable decline in GDP growth of at least 3 or 5 percentage points. They also expand the range of indicators by supplementing five standard macroeconomic variables with two corporate leverage ratios and, more uniquely, a slew of institutional quality variables, including six measures of rule of law, eight of shareholders’ rights, and five of creditors’ rights, each of which is aggregated using principal components analysis. Most importantly, they explore interactions among these variables using a methodology called **binary recursive tree**, which works as follows. First, thresholds for each indicator are identified that minimize (the sum of) Type I and Type II errors. The sample is then split into two “branches” using the threshold of the best indicator. These two steps are then repeated to construct sub-branches, and the process is continued until some stopping rule (which penalizes overfitting via too many branches) is satisfied. In their sample, countries with poor public sector governance are much more likely to have a crisis; and of those with poor public sector governance, those which also have current account deficits greater than 2.6 percent of GDP are more vulnerable than others, and so on. These interactions are not easily captured in the linear structure of standard probit models.

Another study that uses classification rules is Osband and Van Rijckeghem (2000), who identify values for indicators which keep a country safe from currency crises. Using 18 monthly vulnerability indicators for 31 emerging markets over the period 1985-1998, their objective is the **identification of safe or near-safe regions** by the use of three types of filters: simple filters (“ $X > g_1$ ”), intersection filters (“ $X > g_1$ AND $Y > g_2$ ”), and linear combination filters (“ $aX + bY > g_3$ ”). Filters are assessed based on their ability to separate or extract tranquil periods from crisis periods in sample, as well as their “marginal extraction” rate, i.e. their ability to extract tranquil periods which are not extracted by other filters. They find that a set of nine filters—three simple, five intersection and one linear combination—are able to identify 47 percent of tranquil periods as being safe or near-safe. External debt and reserve adequacy feature heavily in these filters. This model, as well as the binary recursive tree model of Ghosh and Ghosh, are useful complements to the standard models. Osband and Van Rijckeghem suggest that their model can be used as the first part of a two-stage early-warning system: to initially identify which countries are safe or near-safe based on these filters, and then to identify the degree of vulnerability using more standard predictive early-warning systems.

The final study that emphasizes interactions among indicators is Apoteker and Barthelemy (2001). Their model is based on the evaluation of five “fundamental balance” charts. These are scatterplots that illustrate certain aspects of a country’s economy; for example, the growth balance scatterplot combines a domestic growth indicator (GDP per capita growth) with an external balance indicator (current account balance). **A genetic algorithm is used to identify quadrants** in the fundamental balance charts which are most associated with four types of crises: transfer crises, liquidity crises, exchange rate crises and cyclical development crises. Apoteker and Barthelemy define exchange rate crises as movements of the real exchange rate by at least 20 percent in one quarter, 30 percent in two quarters, or 40 percent between three and six quarters. The genetic algorithm provides a set of conditions which are associated with crises, and vulnerability is measured by how many of these conditions are satisfied at a given point in time. Precise definitions of the indicators used are absent, and there is no formal testing of the model. Instead, charts are presented indicating Mexico’s vulnerability to a cyclical crisis—but not to an exchange rate crisis—from the second quarter of 1994 onwards. A similar result is found for Thailand in 1997.

III. A MARKOV-SWITCHING APPROACH TO EARLY-WARNING SYSTEMS

Regime switching models have long been a tool available to empirical economists, with early work on these models going back to Quandt (1958), Goldfeld and Quandt (1973), and Hamilton (1990). Applications have only become common in the last decade, however, with the advent of greater computing power. Markov-switching models with constant transition probabilities have been applied to interest rates (Hamilton 1988), the behavior of GNP (Hamilton 1989), stock returns (Cecchetti, Lam and Mark 1990), and floating exchange rates (Engel and Hamilton 1990).¹⁸ One serious limitation of the earlier Markov-switching models, however, was the restriction of constant transition probabilities. The baseline model was thus extended to allow for time-varying transition probabilities, by Lee (1991) and Diebold, Weinbach and Lee (1994) and used to model long swings in the dollar-pound rate, as well as by Filardo (1993, 1994) to analyze business cycle phases.

As mentioned in the introduction, there are two primary motivations for using Markov-switching model with time-varying probabilities in modeling speculative attacks. First, one can avoid the many ad hoc assumption required in the standard models. Even if, as many of the studies claim, their results are robust to these ad hoc assumptions, we believe there is virtue in simplicity. Second, using exchange rates or the index of speculative pressure directly avoids the loss of information that results when these variables are transformed into a binary crisis dummy variable. In particular, exchange rate dynamics may itself be informative about the likelihood of a large speculative attack. A small increase in volatility

¹⁸ A comprehensive review of the applications of Markov-switching models in econometrics can be found in Nelson and Kim (1999).

(e.g., from a widening of an exchange rate band) or small devaluations in the span of a few months might foreshadow a coming currency crisis, but this information remains unutilized (and in fact is erased by the threshold dating process) in the standard approaches. As we will see below, even small changes in exchange rate behavior are utilized in a regime-switching framework as signs of increasing speculative pressure.

There are three disadvantages to using Markov-switching models. The first is computational; Markov-switching models with time-varying probabilities are still not part of the standard econometric software packages. But this drawback has become minor, as more researchers use the methodology and make their code available, and since software programs such as EViews now allow the creation of general log likelihood objects.¹⁹ A second drawback is the difficulty in testing Markov-switching models against the null of no switching, as one encounters problems with unidentified nuisance parameters (the coefficient parameters in the transition probability matrix), as well as with a singular information matrix. Various tests have been suggested, including Davies (1977, 1987), Hansen (1992, 1996), Hamilton (1996), Garcia (1998) and Mariano and Gong (1998) for testing a constant transition probability model against a null of no switching. For the time-varying transition probability case, one can do a sequential test: first, test the constant transition probability model against a null of no switching, and then test the time-varying transition probability model against a constant transition probability model. Note that testing the significance of individual coefficient estimates, as well as testing the overall model against a null of constant switching, can easily be done using standard t -statistics and likelihood ratio tests. The third drawback is that the likelihood surface can have several local maxima and is sometimes ill-behaved. The model may fail to converge when too many explanatory variables are included, and t -statistics may be sensitive to the choice of step size, since derivatives are calculated numerically. Thus, a judicious choice of for start-up values and step size in the maximum likelihood estimation is important.

A. Model Specification and Estimation

The latent variable in the model follows a first-order, two-state Markov chain $\{s_t\}_{t=1}^T$, where $s_t=1$ denotes a crisis state and $s_t=0$ denotes a tranquil state. Although s_t is not directly observable, the behavior of our dependent variable y_t —which can be either the nominal exchange rate change or the speculative pressure index—is dependent on s_t as follows:

$$y_t | s_t \overset{iid}{\sim} N(\mu_{s_t}, \sigma_{s_t}^2) \quad (1)$$

¹⁹ The EViews code and the dataset used for estimating the models in this paper are available from the author on request.

so that both the mean and variance of y_t can shift with the regime.²⁰ The density of y_t conditional on s_t is then

$$f(y_t|s_t) = \frac{1}{\sqrt{2\pi}\sigma_{s_t}} \exp\left(\frac{-(y_t - \mu_{s_t})^2}{2\sigma_{s_t}^2}\right) \quad (2)$$

for $s_t = 0, 1$.

The latent regime-switching variable s_t evolves according to the transition probability matrix \mathbf{P} :

$$\begin{array}{cc} & \text{Time } t \\ & \begin{array}{cc} \text{State 0} & \text{State 1} \end{array} \\ \begin{array}{c} \text{State 0} \\ \text{State 1} \end{array} & \begin{array}{cc} \left[\begin{array}{cc} p'_{00} & p'_{01} = (1 - p'_{00}) \\ \Pr(s_t = 0|s_{t-1} = 0, x_{t-1}) & \Pr(s_t = 1|s_{t-1} = 0, x_{t-1}) \\ = F(x_{t-1}' \beta_0) & = 1 - F(x_{t-1}' \beta_0) \\ \hline p'_{10} = (1 - p'_{11}) & p'_{11} \\ \Pr(s_t = 0|s_{t-1} = 1, x_{t-1}) & \Pr(s_t = 1|s_{t-1} = 1, x_{t-1}) \\ = 1 - F(x_{t-1}' \beta_1) & = F(x_{t-1}' \beta_1) \end{array} \right] \end{array} \end{array} \quad (3)$$

where p'_{ij} is the probability of going from state i in period $t-1$ to state j in period t , and F is a cumulative distribution function, most typically the logistic or the normal c.d.f. The elements of the $k \times 1$ vector x_{t-1} are the early-warning indicators that can affect the transition probabilities.

One final quantity needed to complete the model is the start-up value $p_1^1 = \Pr(s_1 = 1)$, which gives the unconditional probability of being in state 1 at time 1. As Diebold, Weinbach and Lee (1994) note, the treatment of this quantity depends on whether x_t is stationary or not. If x_t is stationary, then p_1^1 is simply the long-run probability that $s_1=1$, which in turn would be a function of (β_0, β_1) . If x_t is nonstationary, then p_1^1 is an additional parameter that must be

²⁰ An autoregressive process for y_t can be assumed as well, and the autoregressive parameters can switch from one regime to another, if desired. The assumption of normality can also be relaxed.

estimated. In practice, for a long enough time series this value has a negligible effect on the likelihood function, and whether one calculates it as a function of (β_0, β_1) , estimates it as a separate parameter, or just sets it at a constant value makes little difference.

The estimation procedure we use is direct maximization of the likelihood, where the likelihood function is calculated using the iteration described in Hamilton (1994, pp. 692-93). Using information available up to time t , we can construct $\Pr(s_t = j | \Omega_t; \theta)$, the conditional (filtered) probability that the t th observation was generated by regime j , for $j = 1, 2, \dots, N$, where N is the number of states (in this paper, $N=2$). Collect these conditional probabilities into an $(N \times 1)$ vector $\hat{\xi}_{t|t}$.

One can also form forecasts using the conditional (forecast) probability of being in regime j at time $t+1$, given information up to time t : $\Pr(s_{t+1} = j | \Omega_t; \theta)$, for $j = 1, 2, \dots, N$. Collect these forecast probabilities in an $(N \times 1)$ vector $\hat{\xi}_{t+1|t}$. Lastly, let η_t denote the $(N \times 1)$ vector whose j th element is the conditional density of y_t in equation (2). These filtered and forecast probabilities are calculated for each date t by iterating on the following equations:

$$\hat{\xi}_{t|t} = \frac{(\hat{\xi}_{t|t-1} \circ \eta_t)}{\mathbf{1}'(\hat{\xi}_{t|t-1} \circ \eta_t)} \quad (4)$$

$$\hat{\xi}_{t+1|t} = \mathbf{P}'_{t+1} \hat{\xi}_{t|t} \quad (5)$$

where \mathbf{P}_t is the $(N \times N)$ transition probability matrix going from period $t-1$ to period t , described in equation (3), and \circ denotes element-by-element multiplication. Equation (4) calculates $\Pr(s_t = j | \Omega_t; \theta)$ as the ratio of the joint distribution $f(y_t, s_t = j | \Omega_t; \theta)$ to the marginal distribution $f(y_t | \Omega_t; \theta)$, the latter being obtained by summing the former over the states $1, 2, \dots, N$. Equation (5) implies that once we have our best guess as to what state we are in today, we just pre-multiply by the transpose of the transition probability matrix \mathbf{P} to obtain the forecast probabilities of being in various states in the next period.

Given an initial value for the parameters, θ , and for $\hat{\xi}_{1|0}$, which in our model is just $[1 - p_1^1, p_1^1]'$, we can then iterate on (4) and (5) to obtain values of $\hat{\xi}_{t|t}$ and $\hat{\xi}_{t+1|t}$ for $t = 1, 2, \dots, T$. The log likelihood function $L(\theta)$ can be computed from these as

$$L(\theta) = \sum_{t=1}^T \log f(y_t | X_t, Y_{t-1}; \theta) \quad (6)$$

where

$$f(y_t | X_t, Y_{t-1}; \theta) = 1'(\hat{\xi}_{t|t} \circ \eta_t). \quad (7)$$

One can then evaluate this at different values of θ to find the maximum likelihood estimate.

IV. DATA DESCRIPTION AND TRANSFORMATION

The model is estimated using monthly data from January 1972 to December 1999 for the five Asian crisis countries: Indonesia, the Republic of Korea, Malaysia, the Philippines, and Thailand. The dependent variable in our model is the month-to-month percentage change in the nominal exchange rate. Nothing precludes the use of the speculative pressure index as the dependent variable, if one is interested in unsuccessful speculative attacks as well. Another alternative to using the index of speculative pressure, which avoids the need to weight the various components, is described in Abiad (2002). That paper adds reserve changes and interest rate changes as dependent variables, in addition to exchange rate changes, but rather than combining the three variables into a weighted average, the variables are stacked into a 3x1 vector whose distribution is dependent on the Markov-switching crisis variable s_t . The main finding is that adding reserve and interest rate changes to the model does not help identify any additional crisis episodes not already picked up by the univariate model based on exchange rates alone.

We explore a broad set of twenty-two early-warning indicators, which are listed in Table 1. The indicators can be classified into three categories. The first group includes standard measures of *macroeconomic imbalance*. There are three measures of external imbalance: deviations of the real exchange rate from a Hodrick-Prescott trend,²¹ the current account balance relative to GDP, and the growth rate of exports. There are three measures of the adequacy of central bank reserves: the level and the growth rate of M2/reserves, and the growth rate of reserves. We also look at credit expansion, as measured by growth rate of real domestic credit. Two measures of real economic activity are used—the growth rate of industrial production and real GDP growth interpolated from quarterly data. Some crises have been preceded by the bursting of an asset market bubble, usually in the equities market or the property market, so we include the six-month change in the country's stock market index as well. Finally, we include the real interest rate.

The second category of indicators relate to *capital flows*. The first indicator in this group is the 3-month LIBOR, which has been a primary determinant of the level of capital flows to emerging markets. A second indicator captures the idea that large capital inflows usually fuel a lending boom; one measure of this lending boom, first used by Sachs, Tornell and Velasco

²¹ Alternative methods of detrending produce similar results.

Table 1. Early-Warning Indicators

| Category | Concept | Measure |
|--------------------------------|---------------------------------------|--|
| Macroeconomic indicators | External imbalance/real overvaluation | 1 Deviations of real exchange rate from trend |
| | | 2 Current account balance/GDP |
| | | 3 Export growth rate |
| | Inadequacy of reserve cover | 4 M2/Reserves, level |
| | | 5 M2/Reserves, growth rate |
| | | 6 Reserves growth rate |
| | Overexpansion of credit | 7 Growth rate of real domestic credit, deflated by nominal GDP |
| | Slowdown in the real economy | 8 Industrial Production, growth rate |
| | | 9 Real GDP, growth rate (interpolated from quarterly GDP) |
| Capital flows indicators | Asset price boom/bust | 10 Stock market performance, growth rate |
| | Monetary tightening | 11 Real interest rate |
| | Possible cause of reversal of flows | 12 LIBOR |
| | Lending boom | 13 Bank assets/GDP, growth rate |
| | Short-term debt | 14 Short-term debt to reserves |
| | Composition of capital flows | 15 Cumulative non-FDI flows/GDP |
| | | 16 Portfolio flows, share in stock of total capital flows |
| Financial fragility indicators | Capital adequacy | 17 Bank reserves/Total bank assets |
| | Bailing out by the central bank | 18 CB credit to banks/Total bank liabilities |
| | Confidence in banks | 19 Bank deposits/M2, level |
| | | 20 Bank deposits/M2, growth rate |
| | Ability of banks to mobilize deposits | 21 Loans/deposits, level |
| | | 22 Loans/deposits, growth rate |

(1996), is the growth in the ratio of bank assets to GDP. The three other indicators in this category focus on the composition of capital flows: the level of short-term debt to reserves, the stock of non-FDI investment (measured as a cumulation of flows) relative to GDP, and the ratio of cumulative portfolio inflows to total cumulative inflows.

The final category includes indicators of *financial fragility*. Kaminsky and Reinhart (1999) have noted that currency crises and banking crises tend to occur together, and that based on their sample of 20 countries over the period 1970–95, problems in the banking sector typically precede a currency crisis. The first indicator of financial sector soundness we use is a rough measure of capital adequacy, the ratio of bank reserves to bank assets. A second indicator is central bank credit to banks, relative to total banking liabilities; an increase in central bank credit may indicate financial weakness, if its purpose is to prop up or bail out weak banks. The ratio of bank deposits to M2 indicates the relative confidence that households and businesses have in the banking system, with a low ratio indicating a lack of confidence; we include both the level and the growth rate of this ratio. Finally, we look at both the level and the growth rate of the loan-deposit ratio. A high and/or rising loans-to-deposits ratio may indicate increased banking system fragility, with an inadequate level of liquidity to respond to shocks. It should also be noted that one of the macroeconomic indicators, the real interest rate, is also frequently used as an indicator of financial sector soundness, as high real interest rates often lead to an increase in nonperforming loans.²²

Given the large number of indicators, a general-to-specific procedure was used to pare down the set and identify the final model. For each country, the model was run using each of the 22 early-warning indicators, one at a time.²³ The coefficient estimates from these regressions can be found in Table 2. The coefficients on the indicators correspond to the parameter β_0 that enters into p'_{00} , the probability of remaining in the tranquil state. All the variables are transformed such that an increase in the variable lowers the probability of remaining in the tranquil state, so that negative coefficient estimate is “correct”.

²² Nonperforming loan ratios were also considered but were unavailable for most countries for a long enough period. Evidence for the Philippines, however, indicates that NPL ratios are a lagging rather than a leading indicator; the NPL ratio actually declined from about 8 percent in 1990 to about 4 percent just before the Asian crisis, and only started increasing substantially in late 1997 and in 1998.

²³ The hill-climbing method using in maximum likelihood estimation will converge very slowly if the various indicators are of very different magnitudes, since a step-size that is fine enough for the small variables will move very slowly for the large ones. To aid in estimation we rescale each variable to be zero mean and unit variance. An alternative transformation, used by KLR and Berg and Pattillo (1999), is to use percentiles.

Table 2. Coefficient Estimates for the 22 EWIs (Bivariate Regressions)

| Indicator | Indonesia | | Korea | | Malaysia | | Philippines | | Thailand | |
|--|-----------|--------|-------|--------|----------|--------|-------------|--------|----------|--------|
| | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat |
| Deviations of real exchange rate from trend | -0.65 | -2.84 | -0.30 | -2.02 | -0.35 | -2.55 | -0.30 | -2.28 | -0.26 | -2.17 |
| Current account balance/GDP | -0.36 | -0.89 | -0.97 | -1.42 | -0.12 | -0.43 | -0.12 | -0.90 | 0.22 | 1.43 |
| Export growth rate | -0.36 | -0.83 | -0.28 | -1.04 | 0.11 | 1.10 | -0.29 | -2.09 | 0.02 | 0.15 |
| M2/Reserves, level | -0.18 | -1.29 | -0.13 | -1.37 | -0.04 | -0.33 | -0.49 | -3.62 | -0.26 | -1.91 |
| M2/Reserves, growth rate | -0.16 | -1.27 | -0.14 | -1.58 | -0.09 | -0.64 | -0.22 | -1.61 | -0.28 | -2.09 |
| Reserves growth rate | -0.37 | -1.52 | 0.03 | 0.23 | 0.04 | 0.29 | -0.28 | -1.45 | -0.45 | -2.13 |
| Growth rate of real domestic credit, deflated by nominal GDP | 0.04 | 0.26 | -0.18 | -1.07 | -0.32 | -2.16 | 0.47 | 1.91 | -0.15 | -0.81 |
| Industrial Production, growth rate | -0.20 | -1.29 | -0.28 | -1.39 | 0.00 | -0.03 | -0.07 | -0.52 | 1.42 | 1.24 |
| Real GDP, growth rate (interpolated from quarterly GDP) | -0.43 | -1.05 | -0.28 | -1.22 | -0.19 | -1.20 | -0.29 | -2.39 | -0.73 | -2.69 |
| Stock market performance, growth rate | -0.12 | -0.06 | -3.14 | -1.27 | -0.24 | -1.01 | 0.32 | 1.83 | -0.13 | -0.89 |
| Real interest rate | -0.17 | -0.57 | 0.12 | 0.87 | -0.15 | -1.08 | 0.11 | 1.06 | -0.39 | -2.21 |
| LIBOR | -0.11 | -0.50 | -0.01 | -0.07 | -0.27 | -2.32 | -0.56 | -3.97 | -0.14 | -1.04 |
| Bank assets/GDP, growth rate | 0.17 | 0.92 | -0.06 | -0.35 | -0.26 | -2.13 | 0.27 | 2.24 | 0.02 | 0.10 |
| Short-term debt to reserves | 0.19 | 0.66 | -0.01 | -0.01 | -0.74 | -2.92 | -1.01 | -3.95 | -0.25 | -0.91 |
| Non-FDI capital flows; stock relative to GDP | 0.26 | 0.81 | -0.01 | -0.02 | -0.01 | -0.07 | -0.70 | -2.41 | -0.43 | -1.65 |
| Portfolio flows, share in stock of total capital flows | -0.22 | -1.24 | -0.40 | -2.39 | 0.10 | 0.55 | -0.29 | -1.86 | -0.35 | -2.00 |
| Bank reserves/Total bank assets | -0.05 | -0.18 | -0.25 | -1.34 | 0.07 | 0.44 | 0.21 | 1.64 | -0.04 | -0.28 |
| CB credit to banks/Total bank liabilities | -0.24 | -1.76 | 0.02 | 0.18 | 0.02 | 0.07 | -0.47 | -3.62 | 0.04 | 0.26 |
| Bank deposits/M2, level | 0.00 | 0.00 | 0.10 | 0.79 | -0.16 | -1.14 | 0.00 | -0.03 | 0.21 | 1.26 |
| Bank deposits/M2, growth rate | -0.23 | -0.58 | 0.13 | 1.01 | -0.04 | -0.32 | -0.10 | -1.01 | 0.13 | 0.65 |
| Loans/Deposits, level | 0.14 | 0.33 | -0.13 | -0.41 | 0.07 | 0.53 | -0.19 | -1.58 | 0.02 | 0.11 |
| Loans/Deposits, growth rate | 0.03 | 0.13 | -0.13 | -0.87 | -0.08 | -0.57 | 0.17 | 1.81 | 0.46 | 1.61 |

Examining the results of Table 2 more closely, we see that the real overvaluation indicator is correctly signed and significant across all five countries. In fact, it is the only indicator that is uniformly correctly signed and significant. Four other variables—the level and growth rate of M2/reserves, the growth rate of real GDP, and the LIBOR—are correctly signed in all cases, but are only occasionally significant. All the other indicators have coefficient estimates that have correct signs and/or are significant for some countries, but not for others. Which brings us to another important point: indicator performance clearly varies widely from country to country. An assumption of parameter equality across countries, underlying EWS model estimates which are based on a panel of countries, may lead to incorrect results, and may contribute to poor predictive performance.

The set of indicators for each country was then narrowed based on which coefficient estimates were correctly signed. Desire to monitor a wider set of indicators suggested against eliminating correctly signed coefficients whose *t*-statistics were not significant at the 5 percent level.²⁴ The moderate correlation among the early warning indicators also suggested that the *t*-statistics may be misleading. In addition, a likelihood-ratio test of the joint significance of the explanatory variables showed them to be significant.

How high must forecast probabilities rise to be warranted as significant? It has been standard practice in the early-warning systems literature to map a model's forecast probability into a binary "alarm signal" by determining some cutoff probability, and letting the signal equal 1 if the forecast probability rose above this threshold, and 0 otherwise.²⁵ An assessment of predictive ability is then conducted by computing the number of crises the model signal correctly calls (by sending an alarm within a particular window, usually 24 months before the actual crisis occurrence), and the number of false alarms the model sends.

²⁴ The more conventional procedure of keeping only correctly signed and statistically significant variables results in only one indicator (real overvaluation) remaining for Indonesia, the Republic of Korea, and Malaysia, and only two indicators remaining for the Philippines (real overvaluation and M2/reserves) and Thailand (real overvaluation and real GDP growth).

²⁵ Although originally applied to individual indicators by Kaminsky, Lizondo, and Reinhart (1997), this methodology has since been used on the overall model by several studies, most notably Berg and others (1999), to evaluate a model's performance, as well as to compare competing early-warning systems.

In this context, it should be noted that forecast probabilities between competing early warning systems are not directly comparable; in particular, one should adjust for the time horizon the model is using. Most of the early warning systems in the literature focus on relatively long-horizon forecasting, with horizons of 12 or 24 months being the norm. The regime-switching model we use here, on the other hand, estimates one-month ahead forecasts. To make forecast probabilities from different models comparable, the forecast horizons must be matched, and the most straightforward way to do this would be to transform the short-horizon forecast into a long-horizon equivalent, using:

$$\begin{aligned}\Pr(\text{crisis over next } n \text{ months}) &= 1 - \Pr(\text{no crisis over next } n \text{ months}) \\ &= 1 - \Pr(\text{no crisis over next 1 month})^n \\ &= 1 - (1 - \Pr(\text{crisis over next 1 month}))^n\end{aligned}\tag{16}$$

Of course, this transformation is made under the assumption that the fundamentals that determine the crisis probability neither worsen nor improve. If the former, then the n -month crisis probability will be higher; if the latter, then the crisis probability will be lower.²⁶ As an example, a 10 percent probability of a crisis over one month would be equivalent, *ceteris paribus*, to a three-month crisis probability of $1-(0.90)^3 = 27$ percent, and a one-year crisis probability of $1-(0.90)^{12} = 72$ percent.

V. ESTIMATION RESULTS

The final model estimates for the five countries can be found in Table 3. For all five countries, State 0 is identified as a low-mean, low-volatility regime while State 1 is a high-mean, high-volatility regime. Average volatility, as measured by the standard deviation, is very low in the tranquil state—less than 1 percent per month in all five countries—while average volatility during crisis periods is quite large, with the highest crisis volatilities estimated for Indonesia, at 29 percent per month. In fact, volatility seems to be the primary distinguishing characteristic between tranquil and crisis periods, as σ_1 is significantly different from σ_0 in all cases. The average depreciation in tranquil periods is effectively zero (less than a quarter percent per month) in all countries, while in crisis periods it ranges from 2.1 percent per month in Thailand, to 12.6 percent per month in Indonesia, but the standard errors are large enough so that the one cannot reject equality of μ_1 and μ_0 . The coefficients on the indicators in the time-varying probabilities are all correctly signed, but as noted earlier, they are insignificant in most cases. This might be due to correlations among the

²⁶ Alternatively, one can construct projected time paths for the early-warning indicators, and use these to calculate an n -month crisis probability. The accuracy will, of course, depend on the reliability of the projected time paths.

Table 3. Final Model Estimates

| Indicator | Indonesia | | Korea | | Malaysia | | Philippines | | Thailand | |
|--|-----------|--------|-------|--------|----------|--------|-------------|--------|----------|--------|
| | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat | Coeff | T-stat |
| Mean, State 0 | 0.20 | 11.13 | 0.15 | 3.09 | 0.01 | 0.14 | 0.05 | 2.44 | 0.00 | 0.09 |
| Mean, State 1 | 12.59 | 1.32 | 2.83 | 0.73 | 4.41 | 1.96 | 2.62 | 2.27 | 2.10 | 1.35 |
| Sigma, State 0 | 0.30 | 27.43 | 0.73 | 31.76 | 0.83 | 21.34 | 0.21 | 21.36 | 0.46 | 27.61 |
| Sigma, State 1 | 29.54 | 8.55 | 12.18 | 7.99 | 4.15 | 4.65 | 6.36 | 13.39 | 8.40 | 10.01 |
| Constant (beta0) | 2.52 | 4.42 | 17.14 | 0.46 | 1.76 | 6.18 | 1.92 | 5.70 | 2.90 | 1.64 |
| Deviations of real exchange rate from trend | -0.60 | -1.00 | -1.13 | -0.41 | -0.21 | -0.93 | -0.47 | -1.91 | -0.10 | -0.27 |
| Current account balance/GDP | | | -4.39 | -0.24 | | | | | | |
| Export growth rate | -0.05 | -0.08 | | | | | -0.14 | -0.32 | | |
| M2/Reserves, level | -0.16 | -0.45 | -1.48 | -0.12 | | | -0.74 | -3.23 | -0.58 | -0.28 |
| M2/Reserves, growth rate | | | | | | | -0.12 | -0.60 | | |
| Reserves growth rate | -0.08 | -0.18 | | | | | | | -0.62 | -0.60 |
| Growth rate of real domestic credit, deflated by nominal GDP | | | | | -0.20 | -0.82 | | | | |
| Industrial Production, growth rate | | | -3.74 | -0.51 | | | -0.31 | -1.14 | -0.21 | -0.12 |
| Real GDP, growth rate (interpolated from quarterly GDP) | | | | | -0.29 | -0.93 | | | | |
| Stock market performance, growth rate | | | -7.25 | -0.69 | | | | | -0.07 | -0.04 |
| Real interest rate | | | | | -0.04 | -0.15 | | | | |
| LIBOR | | | | | -0.14 | -0.69 | | | -1.04 | -0.32 |
| Non-FDI capital flows; stock relative to GDP | | | | | | | | | | |
| Portfolio flows, share in stock of total capital flows | | | -2.35 | -0.36 | | | | | | |
| CB credit to banks/Total bank liabilities | -0.34 | -0.86 | | | | | | | | |
| Bank deposits/M2, level | | | | | -0.26 | -0.47 | | | | |
| Bank deposits/M2, growth rate | -0.29 | -0.51 | | | | | -0.03 | -0.13 | | |
| Constant (beta1) | 0.71 | 2.44 | 0.94 | 3.40 | -0.02 | -0.03 | 0.65 | 3.84 | 0.82 | 3.02 |
| Number of observations | 336 | | 276 | | 288 | | 216 | | 276 | |
| LR Test Statistic for Joint Significance of Indicators | 11.68 | | 17.61 | | 9.17 | | 25.92 | | 21.07 | |
| P-value | 0.07 | | 0.01 | | 0.16 | | 0.00 | | 0.00 | |

indicator variables; in fact, likelihood-ratio tests for the joint significance of the indicators are significant for all countries except Malaysia, where the test of joint significance is marginally insignificant, with a p-value of 0.16. We now turn to a country-by-country analysis.

A. Indonesia

In the estimated model for Indonesia (Table 3), six indicators are used—real overvaluation, export growth, the level of M2/reserves, reserve growth, central bank credit to the banking sector, and growth of the M2/deposits ratio. There are five speculative pressure episodes in Indonesia in our 1972–1999 sample (Table 4)—a devaluation of 50 percent in November 1978, currency volatility in late 1982 that culminated in a 38 percent devaluation in April 1983, moderate volatility and a 5 percent depreciation in mid-1984, a 44 percent devaluation in September 1986, and the Asian crisis which began in July 1997 with a 6 percent decline in the rupiah. Figure 1 plots these crisis dates, along with 12-month forecast probabilities and alarm signals based on a 50-percent cutoff. Alarm signals are sent at least once in the 12 months preceding four of the five crisis episodes, with the only uncalled crisis being the smallest one, the 5 percent depreciation in mid-1984. However, the Asian crisis was not well-signaled for Indonesia; an alarm was generated only in one month (October 1996), and reflected increased currency volatility during that period. The forecast probabilities do increase steadily, to 45 percent in June 1997, but stay below the signaling threshold of 50 percent.

B. Korea

The indicators that enter the final model for Korea are real overvaluation, the current account to GDP ratio, the level of M2/reserves, industrial production growth, stock market performance, and the share of portfolio flows in total capital flows (Table 3). Three crisis periods are included in the sample (Table 5)—a 20 percent devaluation in January 1980, a depreciation of 7 percent in September–November of the same year (which was likely a continuation of earlier speculative pressure), and the Asian crisis (Figure 2).²⁷ Interestingly, the model already identifies March 1997, when the won depreciated by 4 percent, as a period of speculative pressure. In terms of predicting these episodes, the model does not anticipate the January 1980 depreciation; however, after the initial devaluation it continues to send signals in anticipation of further speculative pressure, which did occur later that year.

²⁷ Current account data for Korea only begins in 1977; hence two earlier devaluations of 14 percent in June 1971 and 21 percent in December 1974 are not covered. The small number of crises and the closeness of the fit in Figure 2 raise the concern that the model may be overfitting the data in the Korean case.

Table 4. Speculative Pressure Episodes and Alarm Signals in Indonesia

| Episodes Identified | Description | Signaled/Anticipated by Model? |
|-------------------------|--|--|
| November 1978 | Devaluation of 50 percent | Yes, from February-September 1978 |
| October 1982-April 1983 | Currency volatility; devaluation of 38 percent in April 1983 | Yes, from August 1982 to September 1983 (sporadic signals) |
| August-September 1984 | Devaluation of 4 1/2 percent over three months | No |
| September 1986 | Devaluation of 44 percent | Yes, from February 1985 to June 1986 |
| July 1997 onwards | Asian Crisis | Yes, but only one month (October 1996) due to increased volatility in the rupiah; probabilities rise from 22 percent in November 1996 to 45 percent in June 1997, but do not cross 50 percent signal threshold |

Table 5. Speculative Pressure Episodes and Alarm Signals in the Republic of Korea

| Episodes Identified | Description | Signaled/Anticipated by Model? |
|--------------------------------------|---|--|
| January 1980 | Devaluation of 20 percent | No |
| October 1980 | Depreciation of 7 percent over three months | Yes, from January-May 1980 (i.e., more volatility was expected after January depreciation) |
| March 1997; and October 1997 onwards | Asian Crisis (October 1997 onwards), but model also detects increased volatility in early 1997 and already identifies March depreciation (4 percent) as speculative pressure period | Yes, but only from February 1997 onwards |

Figure 1. Indonesia: Crisis Dates, Forecast Probabilities, and Alarm Signals

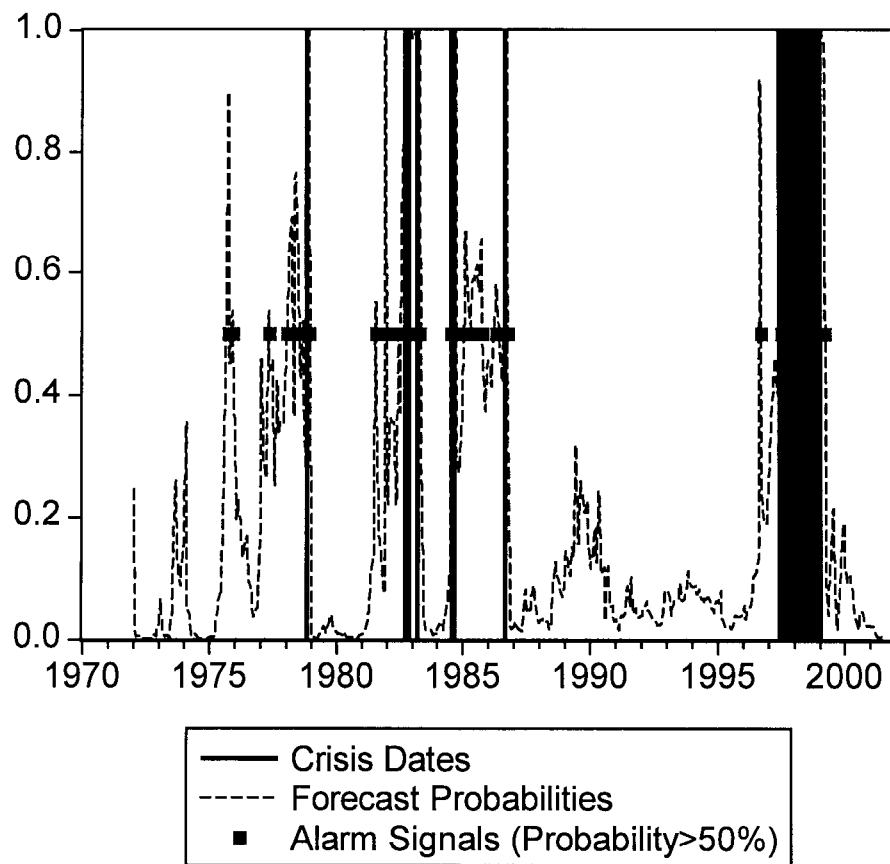
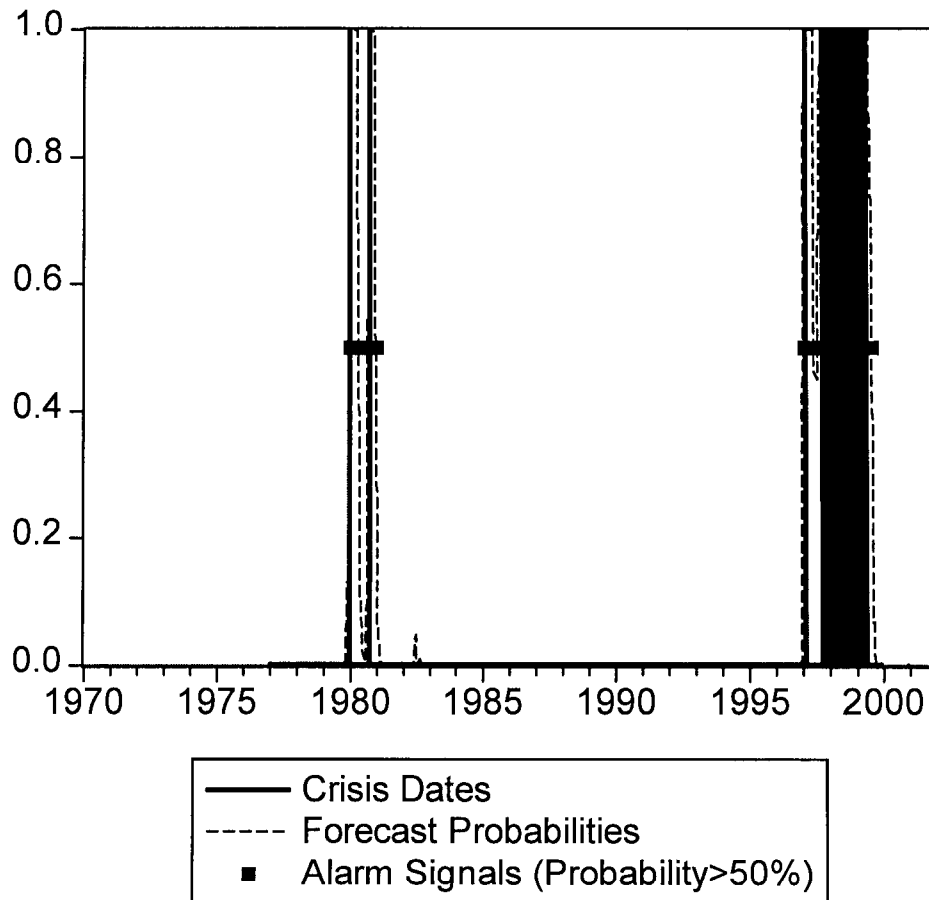


Figure 2. Republic of Korea: Crisis Dates, Forecast Probabilities, and Alarm Signals



With regard to the Asian crisis, the Korean model illustrates the gains from letting an EWS model use information available in the exchange rate behavior. There was already a moderate increase in the won's volatility before the Asian crisis, beginning as early as the middle of 1996, when the won depreciated by 3 percent. As a result of this increased volatility—and combined with Korea's weakening external position, a decline in the stock market and the high share of portfolio flows—the model begins signaling in February 1997.

C. Malaysia

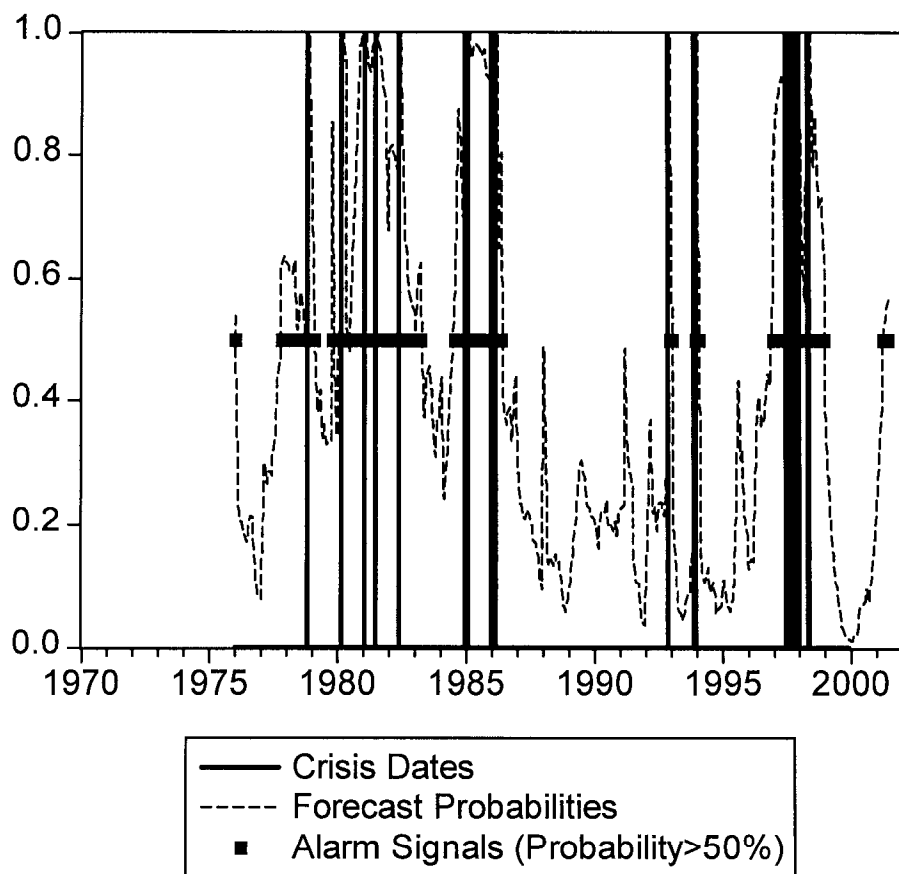
The final model for Malaysia contains six indicators—real overvaluation, domestic credit growth, real GDP growth, the real interest rate, the LIBOR, and the ratio of M2 to deposits. Relative to the four other countries, Malaysia's exchange rate regime was much less of a peg and more of a dirty float. Thus, unlike the other countries which experienced rarer but sharper devaluations, the speculative pressure episodes in Malaysia are protracted periods characterized by increased volatility and a slow deterioration of the exchange rate. Thus, instead of identifying the individual spikes, we group them into four periods which are described in the Table 6.

Table 6. Speculative Pressure Episodes and Alarm Signals in Malaysia

| Episodes Identified | Description | Signaled/Anticipated by Model? |
|----------------------------|---|--------------------------------|
| October 1978-June 1982 | Volatility; 12 percent depreciation over the period | Yes, from November 1977 |
| January 1985-March 1986 | Depreciation of 15 percent | Yes, from August 1984 |
| December 1992-January 1994 | Depreciation of 9 percent | No |
| July 1997 onwards | Asian crisis | Yes, from January 1997 |

The model is able to anticipate three of Malaysia's four speculative pressure periods (Figure 3). Analyzing the individual indicators that enter the model, one finds that the rise in world interest rates contributed to the speculative pressure that occurred in the late 1970s and early 1980s; overvaluation, high real interest rates, and a slowdown in real growth contributed to the 1985–86 depreciation; and overvaluation and a domestic credit boom increased vulnerability in the run-up to the Asian crisis.

Figure 3. Malaysia: Crisis Dates, Forecast Probabilities, and Alarm Signals



D. The Philippines

The six indicators that enter into the Philippine model are real overvaluation, export growth, both the level and the growth rate of M2/reserves, industrial production growth, and the growth rate of deposits/M2. The model was estimated from 1982 onwards, since data on industrial production growth is unavailable before 1982. The model identifies three protracted periods of speculative pressure (Table 7).²⁸ The first is from August 1982 to April 1986, when a financial crisis and political turmoil resulted in high exchange rate volatility and a 140 percent depreciation over the period. The second is from June 1988, a period of moderate volatility where the peso depreciated by 34 percent. The third period is the Asian crisis, which began for the Philippines with a 10 percent depreciation in July 1997.

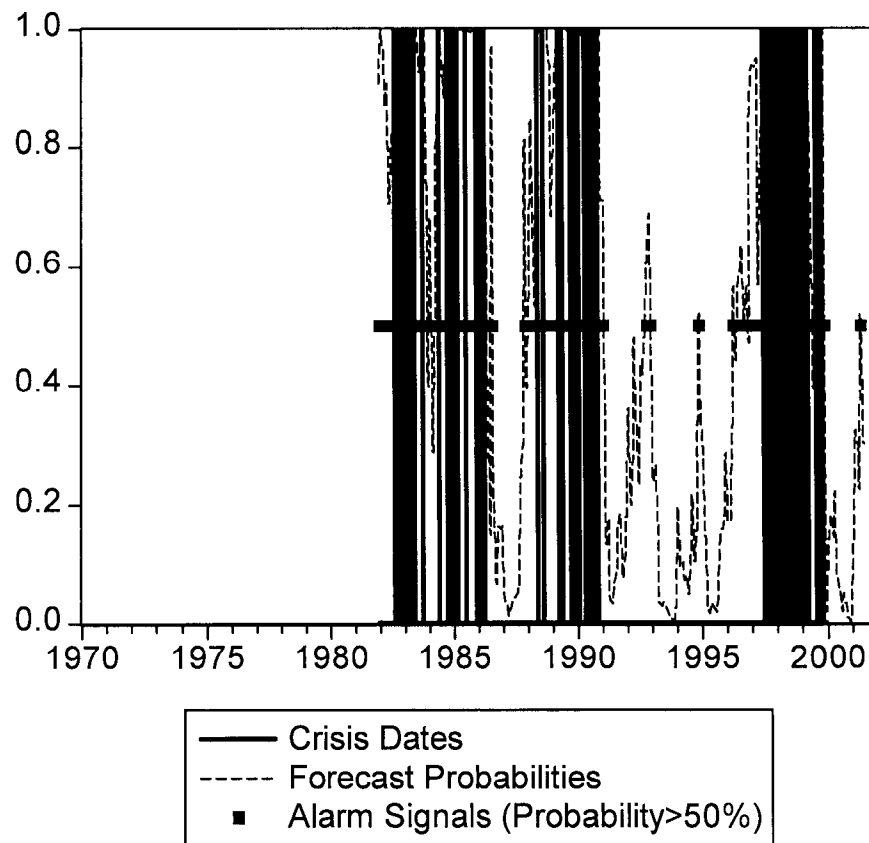
Table 7. Speculative Pressure Episodes and Alarm Signals in the Philippines

| Episodes Identified | Description | Signaled/Anticipated by Model? |
|-------------------------|--|--------------------------------|
| August 1982-April 1986 | Depreciation of 140 percent; high volatility (s.d. 7 percent) | Yes, from January 1982 |
| June 1988–November 1990 | Depreciation of 34 percent; moderate volatility (s.d. 2 percent) | Yes, from December 1987 |
| July 1997 onwards | Asian Crisis | Yes, from May 1996 |

The model for the Philippines is able to anticipate these three crisis periods (Figure 4). Analyzing the individual indicators, one finds that different factors were behind each crisis. A slowdown in both exports and industrial production played some role in triggering the crisis in the early 1980s, but a rise in both the level and growth rate of M2/reserves, as well as a sharp fall in the deposits/M2 ratio (an indicator of the banking crisis that occurred), played a role in prolonging the crisis. Reserve adequacy, as measured by both the level and growth rate of M2/reserves, also increased vulnerability in the late 1980s, which culminated in a 17 percent depreciation in the latter half of 1990, during the Gulf War. Finally,

²⁸ Note that the period in the early to mid 1990s, when the peso was allowed to float more freely, was omitted so as not to be identified as a crisis period.

Figure 4. Philippines: Crisis Dates, Forecast Probabilities, and Alarm Signals



weakening competitiveness that began in late 1996—resulting from the appreciation of the yen against the dollar, to which the peso was pegged—increased the Philippines’ vulnerability, and resulted in the depreciation of the peso following the float of the Thai baht in July 1997.

E. Thailand

The final model for Thailand includes the following indicators: real overvaluation, the level of M2/reserves, reserve growth, real GDP growth, the real interest rate, and the share of non-FDI capital flows in total flows. There are three crisis periods in the sample, a 10 percent depreciation in 1981, a 19 percent depreciation between November 1984–December 1985, and the Asian crisis which began in Thailand in July 1997 (Table 8).

Table 8. Speculative Pressure Episodes and Alarm Signals in Thailand

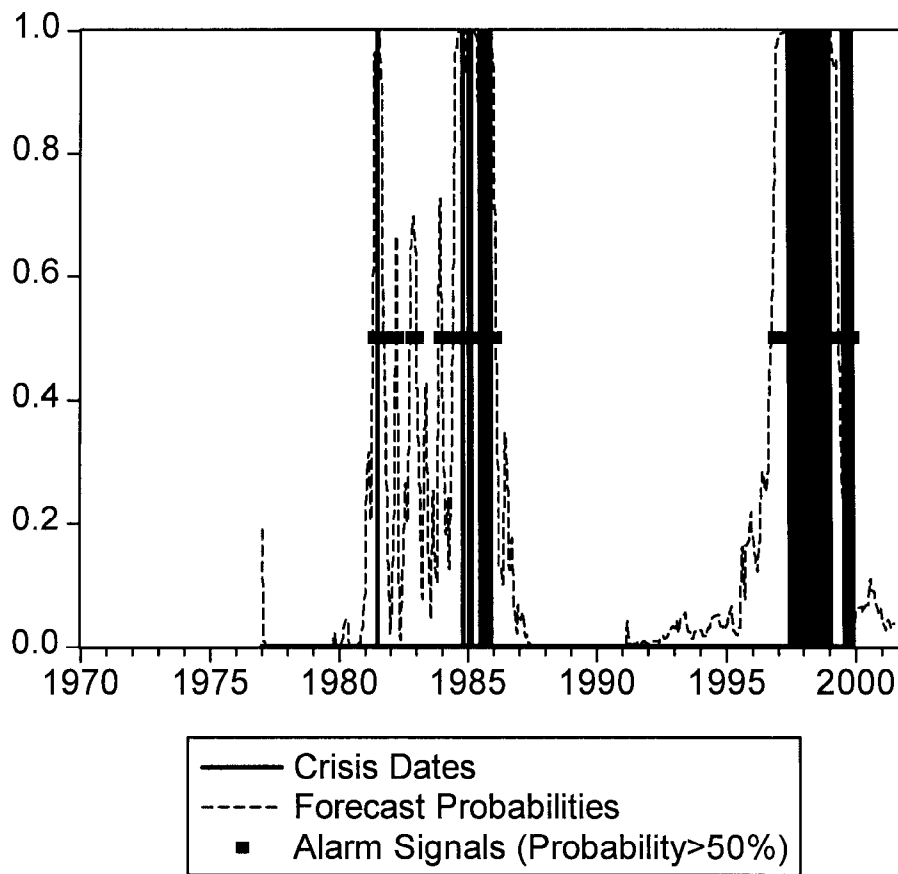
| Episodes Identified | Description | Signaled/Anticipated by Model? |
|-----------------------------|----------------------------|---|
| July 1981 | Devaluation of 10 percent | Yes, but only two months ahead (May 1981) |
| November 1984–December 1985 | Depreciation of 19 percent | Yes, in December 1983–January 1984 and from July 1984 |
| July 1997 onwards | Asian Crisis | Yes, from December 1996 |

All three crisis periods are anticipated, as can be seen in Figure 5. However, signals are sent only two months prior to the July 1981 devaluation. Better warning is provided for the latter crisis episodes. Real overvaluation seems to have played some role in all three crises. Reserve inadequacy, as measured by the level of M2/reserves, also played a role in the 1980s episodes, but not in the Asian crisis. Based on the model, three factors seem to have played a role in increasing Thailand’s vulnerability to crisis in 1997—a loss of external competitiveness, a slowdown in the real economy, and an increasing proportion of non-FDI flows in total capital flows.

VI. FORECAST ASSESSMENT

We now perform a more rigorous evaluation of the predictive performance of the model, both in-sample and out-of-sample. Table 9 contains in sample goodness-of-fit tables for each country model, as well as overall for all five countries. Each 2x2 matrix shows the number of

Figure 5. Thailand: Crisis Dates, Forecast Probabilities, and Alarm Signals



correctly called tranquil and crisis periods, as well as the number of false alarms and the number of missed signals. We summarize this information further in Table 10, which also provides goodness-of-fit measures for five other models evaluated in Berg and Pattillo (1999, henceforth BP). The comparison is only meant to be indicative, as the Markov-switching model differs from the five other models in several important ways, beyond just the differences in model specification. First, the identified crisis dates are different. Second, the forecast horizons are not the same—the models reviewed by BP all use a forecast horizon of 24 months, as opposed to the 12-month forecast horizon used here. Third, the data underlying the estimates, and the transformations applied to them, are similar but not identical. Finally, the models in this paper are estimated country-by-country, whereas all five BP models were estimated on a panel of countries, a point we return to below.

We see that Markov-switching model correctly calls 81 percent of observations. This is slightly lower than the 82–85 percent performance of the standard models in BP, when they use a 50 percent cutoff probability. However, the high predictive performance in those models is driven mostly by their ability to call tranquil periods correctly; they correctly classify 98–100 percent of tranquil periods, as opposed to 89 percent in the Markov-switching model. But in terms of correctly called crises, the Markov-switching model performs much better, calling 65 percent of pre-crisis periods correctly—that is, sending signals in 65 percent of the months where a crisis ensued within a year’s time. The standard models, in contrast, only call 7–19 percent of pre-crisis months correctly. The poorer performance is probably due in part to the longer 24-month forecast horizon they aim for, and also because the 50 percent signaling threshold is too high for those models. BP also report goodness-of-fit for the standard models when the cut-off is lowered to 25 percent, which we replicate in the bottom half of Table 10. The lower signaling threshold increases the number of correctly called pre-crisis periods—now the models correctly send signals in 41–48 percent of pre-crisis periods—but this comes at the expense of a much higher fraction of false alarms. With the lower threshold, false alarms account for 57–65 percent of total alarms, i.e., almost two out of every three signals is a false alarm.

Although the Markov-switching specification probably accounts for part of the improved performance, it is also possible that a substantial portion of the improved performance is due to the fact that the standard models estimate the data using a panel of countries, and assuming that the coefficients are uniform across countries. As we saw in Section V, indicators that matter for crises in one country may not even be pointing in the right direction during crises in another country.

Table 9. Forecast Assessment

| | | Crisis within 12 mos. | No crisis within 12 mos. |
|-------------|------------|-----------------------------|--------------------------------|
| Overall | Signal | 270 | 102 |
| | Non-signal | 147 | 812 |
| Indonesia | Signal | 38 | 18 |
| | Non-signal | 47 | 221 |
| Korea | Signal | 26 | 4 |
| | Non-signal | 29 | 205 |
| Malaysia | Signal | 77 | 29 |
| | Non-signal | 48 | 121 |
| Philippines | Signal | 83 | 40 |
| | Non-signal | 2 | 79 |
| Thailand | Signal | 46 | 11 |
| | Non-signal | 21 | 186 |

What are the out-of-sample predictions of the country Markov-switching models? The models were estimated using data up to the end of 1999. An attempt to estimate the model up to end-1996, to see whether the Asian crisis was forecastable using the model, was not possible in this case, mainly because the model was estimated country-by-country; eliminating the Asian crisis not only removes the most informative episode in the sample, but also results in overfitting and/or nonconvergence of the maximum likelihood algorithm. Hence the only alternative is to look at model forecasts beyond the end of 1999. Admittedly, the hold-out sample from January 2000–July 2001 is relatively small, and moreover, none of the five countries had a crisis during this period. But it is still an informative exercise to see what kinds of probabilities and signals the country models send.

The forecast probabilities and alarm signals for the out-of-sample period of January 2000–July 2001 can be seen in Figures 1–5. For three of the countries—Indonesia, Korea and Thailand—no alarm signals are sent during the period. There was still a moderate probability (about 20 percent) of a crisis in Indonesia through much of 2000, but vulnerabilities (at least those measured by the indicators in the model) have dropped since then. Thailand has shown lower susceptibility to a crisis, and Korea even less so.

In contrast, the models for Malaysia and the Philippines did signal some vulnerability in the out-of-sample period, although only for a few months. Crisis probabilities in Malaysia were actually dropping toward the end of the estimation period (1999) and were low through most of 2000, but started increasing in the last quarter of 2000 and accelerating in 2001 up until July 2001, the last available data point. In fact, probabilities were high enough that the model began sending signals in May 2001. What was driving this increase in vulnerability? An analysis of the indicators entering the Malaysian model identifies several weaknesses. First, there was a steady decline in competitiveness, as measured by the real exchange rate, through 2000 and 2001.²⁹ Second, there was a slowdown in the real economy. And third, there was an sharp rise in real domestic credit growth.

The Philippines also showed some weaknesses in the out-of-sample period, according to the model. Crisis probabilities were actually low in 2000, but the model starting indicating moderate vulnerabilities beginning in the second quarter of 2001. A spike in the crisis probability led to a signal in June 2001, but probabilities decreased in July, the last data point. There was one primary factor behind the increased vulnerability in the Philippines: a weakened external position, seen most clearly in rapidly contracting exports. Over the January 2000–July 2001 out-of-sample period, then, signals were sent for only 4 out of 95 months: May-July 2001 for Malaysia, and June 2001 for the Philippines, and these reflected vulnerabilities due to external weaknesses present in these two countries at that time.

²⁹ This was also evident in another indicator that does not enter into the Malaysia model, export growth.

VII. CONCLUSIONS

There is a general consensus among economists that early-warning systems, no matter how sophisticated, will not be able to forecast crises with a high degree of accuracy. Even economists who construct such models are aware of this, and see these models as no more than useful supplements to more informed country analyses, and as a means of summarizing information in an unbiased, objective manner. Nevertheless, increased emphasis on crisis prevention (as opposed to crisis resolution) means that policymakers need to utilize all the tools available for assessing countries' vulnerabilities, and to improve these tools when possible. This paper hopes to assist in this effort, first by surveying the recent empirical literature on currency crises, and by analyzing an alternative EWS approach that addresses some of the shortcomings of existing models.

The survey of 30 selected empirical studies written since 1998 is meant to increase awareness of the various econometric approaches to early warning systems that have been developed, so that practitioners have at their disposal a larger set of tools in assessing vulnerability. Many of the proposed approaches look promising, and virtually all report some improvement over the standard probit/logit and indicators models. However, many of the studies do not perform rigorous evaluations of performance. Adoption of standard evaluation procedures—including goodness-of-fit tables and measures, accuracy scores, and out-of-sample testing—will help potential users gauge how useful these models really are. Furthermore, it is difficult to assess relative performance across models, given differences in the datasets used and in the sample of countries studied. A true “horse race” among competing models—where each specification is estimated using the same data and sample of countries—will help resolve this issue.

But even in-sample and out-of-sample tests are only indicative; the true test of these models is in operationalizing them. In this regard, an additional measure of a model's usefulness is simplicity of application. Early warning systems should be easy to replicate and estimate. That is, there should be minimal reliance on ad hoc assumptions, the data should come from published sources, and one should ideally be able to estimate the model using standard software packages or with programming code provided by the authors. If these conditions are satisfied, then it should be possible to monitor these models in real-time at low cost.

In addition to surveying the recent literature on early warning systems, this paper also contributes to it by suggesting an alternative EWS approach based on a Markov-switching model with time-varying transition probabilities. The model does an adequate job of anticipating crises. It correctly anticipates two-thirds of crisis periods in sample (compared to about 50 percent for the standard models), and just as important, sends a much smaller proportion of false alarms. In the January 2000–July 2001 out-of-sample period, no warning signals are sent for three of the five countries studied (Korea, Thailand and Indonesia), but vulnerabilities were signaled for Malaysia and the Philippines in mid-2001, mainly due to a decline in competitiveness and a slowdown in exports.

Beyond the performance of the Markov-switching model itself, there are some lessons that apply to the construction of early warning systems in general. First, accounting for dynamics is important. There is useful information in both the level and the volatility of the exchange rate itself that existing models have ignored. More specifically, some crises have been preceded by a series of smaller depreciations, by a widening of an exchange rate band, and/or an increase in the volatility of exchange rates, and this has not been utilized in existing models. Second, although there are some indicators which are common across countries in their predictive ability (with the real exchange rate being the most uniformly successful), the country-by-country analysis in this paper shows that the performance of individual indicators varies greatly across countries, so that different sets of variables are relevant for different countries. In this light, the one-size-fits-all, panel data approach used in estimating most early warning systems might be one of the causes for their only moderate success. The performance of early warning systems, regardless of the econometric specification chosen, might be improved markedly by taking more care in verifying that the countries used in the estimation possess similar characteristics, or failing that, by estimating the models on a country-by-country basis.

The model presented here is only the simplest variant of what can be done in a Markov-switching EWS. Most obviously, those with a better knowledge of each country can estimate these models using a more informed selection of indicators. Given the role that politics and political stability have played in triggering or exacerbating several crises, most notably Indonesia in 1997, the use of sociopolitical variables could be explored. In light of increased financial globalization, other external factors in addition to world interest rates might be considered, such as global equity market volatility or the spread on high-yield bonds. Regarding the specification itself, one can extend the model in several directions. First, because the focus was on crisis anticipation, the early warning indicators in the current model only affected the probability of moving from a tranquil to a crisis state. But one could let these same indicators (or a different set of indicators) also affect the probability of getting *out* of a crisis. Second, the current model has only two states: a tranquil state, and a speculative pressure state whose main characteristic is high exchange rate volatility. One could extend the model to allow for three (or more) states, where the three states might correspond to tranquil periods, periods of depreciation pressure and periods of *appreciation* pressure. Finally, the issue of modeling contagion across countries within the context of a Markov-switching EWS awaits further investigation.

A Summary of the Recent Empirical Literature on Currency Crises

| | | | | |
|-------------------|---|--|--|--|
| Study | Apoteker and Barthelemy (2001); Apoteker (2000) | Aziz, Caramazza, and Salgado (2000) | Berg and Pattillo (1999b) | Bruggemann and Linne (2000) |
| Dataset | 40 developing countries, monthly and quarterly data from 1970-2001. (p. 5; 2000, p. 19) | 20 industrial and 30 developing countries, monthly and annual data from 1975-1997 (p. 7) | 20 developing and industrial countries, monthly data from 1970-1995 (p. 565) | 6 EU accession candidates, plus Russia and Turkey, monthly data from 1993-2001 (p. 5) |
| Crisis Definition | Exchange rate crisis defined as 20% change in real exchange rate in one quarter, 30% in two quarters or 40% in 3-6 quarters. (p. 8) | Weighted average of exchange rate and reserve changes; threshold is 1.5 times the <i>pooled</i> standard deviation plus the pooled mean; separate calculations were done for high inflation (>80%) countries. (p. 7) | Weighted average of exchange rate and reserve changes; threshold is 3 country-specific standard deviations above the mean; separate calculations for high inflation (>150% six-month inflation) countries. | 20% depreciation against US dollar within ten trading days. (p. 6) |
| Indicators | Eleven variables, combined to measure "growth balance, financing balance, foreign exchange balance, cyclical balance and banking system balance" (p. 7-8) | 23 variables, measuring overheating, external imbalance; also the unemployment rate, short-term capital inflows, and the world interest rate | 15 variables used in KLR (1998), plus the level of M2/reserves, and the current account to GDP ratio. (p. 567) | 16 variables, including capital flight (as measured by BIS deposits); short-term debt; ratio of lending to deposit rates, and bank deposits/GDP (p. 3-5) |
| Approach | Variables are interacted, and quadrants which increase vulnerability are identified using a genetic algorithm | Comparison of pre-and post-crisis behavior of indicators, to their behavior in tranquil periods (p. 11) | Replication/update of KLR results, and three probits using KLR signals, linear, and piecewise linear variables | Composite indicator approach of KLR (1998) |
| Results | No formal testing; charts indicate anticipation of cyclical crises (but not exchange rate crises) in Mexico in 1994, and Thailand in 1997 (p. 11-12) | Useful indicators include overvaluation, terms of trade, inflation, domestic credit growth, M2/reserves, world real interest rates, current account (p. 24-25) | Using the same indicators as KLR, BP find that a linear probit model does better than KLR in predicting the Asian crisis (p. 584) | In addition to overvaluation, weak exports and falling reserves, banking sector indicators were also useful. |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Burkart and Coudert (2000) | Bussiere and Mulder (1999a) | Bussiere and Mulder (1999b) | Caramazza, Ricci, and Salgado (2000) |
|-------------------|---|--|---|--|
| Dataset | 15 emerging market economies, quarterly data from 1980-1998. (p. 1,12) | 23 countries (22 emerging plus Hong Kong), Mexican and Asian crises (46 observations). (p. 9) | 23 countries (22 emerging plus Hong Kong), Mexican and Asian crises (46 observations). (p. 7) | 41 emerging and 20 industrial economies; monthly data from 1990-1998 (p. 7) |
| Crisis Definition | Combination of EMPI threshold, Milesi-Ferretti and Razin (1998) threshold, and amendments "in light of expert judgement." (p. 11) | Weighted average of exchange rate and reserve changes; continuous index is used as dependent variable (p. 12) | Weighted average of exchange rate and reserve changes; continuous index is used as dependent variable (p. 6) | Weighted average of exchange rate and reserve changes; threshold is 1.645 times the <i>pooled</i> standard deviation plus the pooled mean; high (>100%) inflation periods excluded. (p. 7-8) |
| Indicators | An initial set of 34 indicators, pared down to six: reserves/M2, reserves/debt, short-term/total debt, overvaluation, contagion indicator, and inflation (p. 17-20) | 4 political instability measures (p. 6-8), added to the economic variables (RER, lending boom, M2/reserves and dummies) used in Tornell (1999) | Variables in STV (1996), Berg and Pattillo (1999), plus an IMF program dummy; and also many alternative indicators (p. 12-17) | 7 economic indicators plus 6 measures of financial linkages/weakness. (p. 17-18) |
| Approach | Fisher discriminant analysis | OLS of crisis index on indicators | OLS of crisis index on indicators; also some nonlinear interactions in robustness checks (p. 17) | Panel probit for Mexican, Asian and Russian crises |
| Results | Four out of five crises are correctly called, and only one out of five tranquil periods are misclassified as false alarms. | Political instability affects crisis depth; post-electoral periods even more than pre-election periods. | IMF programs significantly reduce crisis depth; liquidity (as measured by reserves/short-term debt) can offset weak fundamentals. | Common creditor increases likelihood of contagion; domestic imbalances important when combined with overvaluation; trade spillovers relevant when CA is weak. |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Cerra and Saxena (2000) | Collins (2001) | Edison (2000) | Eliasson and Kreuter (2001) |
|-------------------|--|--|--|--|
| Dataset | Indonesia, monthly data from 1985-1997 | 25 developing and industrial countries, monthly data from 1985-1998 (p. 13) | 20 developing and industrial countries, monthly data from 1970-1998 (p. 8-9) | 10 emerging markets, monthly data from 1990-2000 (p. 6, 8) |
| Crisis Definition | Determined endogenously in Markov-switching model | Weighted average of exchange rate and reserve changes; threshold is 3 country-specific standard deviations above the mean; separate calculations for high inflation (>150% six-month inflation) countries. (p. 12) | Weighted average of exchange rate and reserve changes; threshold is 2.5 country-specific standard deviations above the mean; separate calculations for high inflation (>150% six-month inflation) countries. (p. 11) | Continuous index constructed in a five-step procedure, fitting a Gaussian followed by an extreme-value distribution (p. 6-7) |
| Indicators | 8 variables in three categories: financial (5), non-financial (2) and political (1). (p. 7-8) | 5 indicators: real overvaluation, current account/GDP, S-T debt/reserves, reserve growth, and export growth (p. 14) | 14 variables, covering the current account, capital account, real sector and financial sector (p. 8) | 12 variables (p. 5) |
| Approach | Markov-switching, with contagion variables (crisis in other countries) entering transition probabilities | Threshold model with latent variable following Brownian motion; distance and drift are functions of indicators | replication and updating of KLR (1998) results; application as single-country EWS (Mexico), and for cross-country comparisons | Multinomial logit with continuous dependent variable, with separate panels for Latin America and Asia |
| Results | Crises in Korea and Thailand significantly increased the likelihood of crisis in Indonesia | Estimates a threshold model, a Poisson model and a probit model and finds that the threshold model fits better than the two alternatives | The model helps identify vulnerabilities, but also gives rise to many false alarms; performance is robust to various sensitivity tests | The continuous crisis variable is more informative and better describes crisis periods (p. 15) |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Esquivel and Larrain (1998) | Ghosh and Ghosh (2002) | Hawkins and Klau (2000) | Herrera and Garcia (1999) |
|-------------------|--|--|--|---|
| Dataset | 15 high-income and 15 middle-income countries, annual data from 1975-1996 (p. 11) | 42 industrial and emerging countries, annual data from 1987-1999 (p. 11) | 24 emerging markets, quarterly data from 1993-1998 (p. 14) | 8 Latin American countries, monthly data from 1980-1998 (p. 3 and Table 1) |
| Crisis Definition | Change in the <i>real</i> exchange rate > 15%, or greater than 2.54 standard deviations that is also greater than 4%. (p. 9) | "Deep" currency crises; those where wtd. ave. of exchange rate and reserve changes is more than 2 country-specific standard deviations above the mean, AND the GDP growth rate declines by at least 3 (or 5) percentage pts. | Scores from -2 to 2 are assigned based on reserve changes, <i>real</i> interest rate changes, and exchange rate changes at both the 3-month and 12-month horizons. (p. 26) | Unweighted sum of exchange rate, reserve and interest rate changes; threshold is 1.5 standard deviations above the mean; separate calculations for hyperinflation episodes (p. 2-3) |
| Indicators | 7 standard indicators (p. 14-16, 22) | 6 rule of law, 8 shareholders rights and 5 creditor rights variables (aggregated by principal components), along with 5 standard indicators and 2 corporate leverage ratios. (p. 11) | Six external vulnerability indicators and five banking system indicators, scored from 0 to 10. (p. 27-28) | 5 indicators, normalized and then summed into a single indicator from which signals are extracted. (p. 3, 11) |
| Approach | Panel probit with random effects | Binary recursive tree, which identifies thresholds, splits sample into branches based on best indicator, and continues branching until stopping rule is reached | Summary tables which present vulnerability scores; assessment of the score is done via a panel probit with fixed effects (p. 19) | Indicators are normalized and summed into a single indicator, from which signals are extracted; detrending via H-P filter, MA and ARIMA also considered (p. 4) |
| Results | All seven indicators, including a regional contagion variable, are significant in explaining occurrence of crisis. | Macroeconomic imbalances are often the trigger for crises, but interactions with institutional variables are important as well | Objective is to present vulnerabilities in a simple, transparent manner, by scoring variables on a 0-10 scale. | The aggregate indicator signals periods of vulnerability. Type I and Type II errors are still high, although lower than in other studies. |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Grier and Grier (2001) | Kamin, Schindler, and Samuel (2001) | Kaufmann, Mehrez, and Schmukler (2000) | Krkoska (2001) |
|-------------------|---|---|--|---|
| Dataset | 25 developing countries, 1997 cross-section (p. 139) | 26 emerging market economies, annual data from 1981-1999 (p. 3) | 58 countries, annual survey results in 1996-1998 (p. 12) | 4 transition countries, quarterly data from 1994-1999 (p. 47) |
| Crisis Definition | Exchange rate depreciation, and stock market returns (p. 142, 144) | Weighted average of two-month changes <i>real</i> exchange rate and reserves; threshold is 1.75 country-specific standard deviations above the mean; no separate calculations for high inflation countries. (p. 4) | Standard deviation of monthly exchange rate changes | Unweighted average of exchange rate, reserve and interest rate changes; two alternative thresholds, t(5%) and t(0.5%) times the standard deviation over the past 36 months. (p. 43) |
| Indicators | Exchange rate regime; controls include 1996 depreciation, current account, external debt/GDP, M2/reserves, lending boom, and real appreciation (p. 141-143) | 6 domestic variables, 4 external balance variables, and 3 external shock variables. (p. 8-9) | Private information of managers; 12 other variables used as controls (p. 28-29) | 10 variables, most standard; one less conventional variable is the CA-FDI gap (p. 44) |
| Approach | OLS regression to investigate whether exchange rate regime affects the degree of depreciation, and stock market returns (p. 143, 146) | Probit; estimate relative contributions of domestic, external balance and external shock variables | Private information of managers is extracted using ordered probit, and used with 12 other variables in OLS regression to explain ER volatility | VAR with ISP(t) as one component, with many ad hoc restrictions imposed. (p. 46-48) |
| Results | Countries with a peg depreciated more, and had lower stock returns. | Overall, domestic factors have a larger contribution to crisis probabilities; but when looking at crisis years only, contribution of external shocks is large; interpretation is that external shocks push vulnerable countries "over edge" | Private information of local managers is significant predictor of exchange rate volatility. | The most significant predictor of crisis vulnerability is the CA-FDI gap; when this exceeded 5% of GDP, a crisis always followed. (p. 54) |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Kumar, Moorthy, and Perraudin (2002) | Kwack (2000) | Martinez Peria (2002) | Mulder, Perrelli, and Rocha (2002) |
|-------------------|---|---|---|--|
| Dataset | 32 emerging market economies, monthly data from 1985-1999 (p. 3, 8) | 7 Asian countries, annual data from 1995-1997 (14 obs. in regression) | 7 ERM member countries, monthly data from 1979-1993 (p. 302) | 19 emerging market countries, monthly data from 1991-1999 (p. 8-9) |
| Crisis Definition | 5% or 10% depreciation, after adjusting for the interest rate differential ("unanticipated crash"), or which is a doubling of the previous depreciation ("total crash"). (p. 6) | Weighted average of exchange rate and reserve changes; continuous index is used as dependent variable (p. 196) | Determined endogenously in Markov-switching model | Berg and Pattillo (1999) crisis definition, and continuous crisis index used in Bussiere and Mulder (1999). (p. 7, 14) |
| Indicators | 17 variables, covering standard macroeconomic variables, financial flow variables, contagion, and other external factors (p. 8-9) | LIBOR; NPL ratio; CA/GDP, S/T debt/total debt, and credit to the private sector. (p. 197) | 6 indicators (domestic credit growth, trade balance, real exchange rate, interest rate differentials, unemployment rate and government surplus (p. 308) | Corporate and macroeconomic balance sheet variables, and legal indicators (in addition to the standard EWS variables) (p. 5-7) |
| Approach | Panel logit | OLS of crisis index on indicators; also OLS of NPL ratio on its determinants (p. 197-198) | Markov-switching | Berg and Pattillo (1999) panel probit, and Bussiere and Mulder (1999a) cross-sectional regression |
| Results | Good performance both in and out of sample; evaluation based on probability timeplots, goodness-of-fit tables, accuracy scores, and profitability of trading strategy. | LIBOR and the NPL ratio (and the debt-equity ratio of firms, which explains the NPL ratio) were the main factors behind the Asian crisis. | The regime-switching framework improves the ability to identify speculative attacks. Budget deficits and expectations affect crisis probabilities. | High leverage and short maturity structures increase both the likelihood and depth of crises. Shareholder rights also have a large impact on crisis probabilities. |

A Summary of the Recent Empirical Literature on Currency Crises

| | | | |
|-------------------|--|--|---|
| Study | Nag and Mitra (1999) | Nitithanprapas and Willett (2000) | Osband and Van Rijckeghem (2000) |
| Dataset | Indonesia, Malaysia and Thailand, monthly data from 1980-1998 (p. 189) | 26 emerging markets, for both the Mexican and Asian crises (52 observations). (p. 19-20) | 31 emerging markets, monthly data from 1985-1998 (p. 244) |
| Crisis Definition | Weighted average of exchange rate and reserve changes; threshold is 2 country-specific standard deviations above the mean. (p. 185) | Weighted average of exchange rate and reserve changes; continuous index is used as dependent variable (p. 8-13) | Exchange rate change > 10%, AND greater than mean plus two standard deviations (mean calculated over the past year, s.d. over the past two years). (p. 244) |
| Indicators | 16 variables in indicators approach (p. 189), 10-12 indicators plus lags (13-23 total variables per country) for the ANN models (p. 208-210) | Interactions between fiscal deficit and real exchange rate, FDI and the current account; controls include private sector credit, M2 and S/T debt to reserves. (p. 13-23) | 18 variables (p. 243) |
| Approach | Artificial neural network; comparison to country-by-country indicators approach | OLS slope dummies regression. | Identifies indicator ranges (half-planes or quadrants) where crises have not occurred. |
| Results | Out-of-sample crisis forecasts of around 80 percent before the Asian crisis. | Interactions between the real exchange rate, FDI and the current account are significant, as are lending booms; adequate reserves reduce risks. (p. 23-35) | Various filters, especially those featuring external debt and reserve adequacy, help identify 47 percent of tranquil periods as being "safe or near-safe". |

A Summary of the Recent Empirical Literature on Currency Crises

| Study | Vlaar (2000) | Weller (2001) | Zhang (2001) |
|-------------------|---|---|---|
| Dataset | 31 emerging markets, monthly data from 1987-1996 (p. 12) | 26 emerging economies, monthly data from 1973-1998 (p. 106-107) | Indonesia, Korea, Philippines and Thailand, monthly data from 1993-1997 (p. 11) |
| Crisis Definition | Weighted average of exchange rate and reserve changes; continuous index is used as dependent variable, but in model evaluation, a threshold of 10 is used to define crisis periods (p. 13) | Weighted average of exchange rate and reserve changes; threshold is 3 country-specific standard deviations above the mean. (p. 118) | Separate thresholds for exchange rate and reserve changes; threshold is 3 standard deviations above mean, where s.d. is calculated using 3-year moving window. (p. 6) |
| Indicators | 14 variables, some of which affect the crisis index directly, others affecting the volatility, and still others the weighting between crisis and tranquil distributions. (p. 11-12) | 16 variables (p. 112) | 4 variables (real exchange rate, inflation, level and growth rate of M2/reserves), as well as contagion measure. (p. 12, 14) |
| Approach | Crisis index is drawn from mixture of normal distributions, where means, volatilities and relative weights are functions of indicators. (p. 11) | Separate logit regressions on pre- and post-liberalization samples | Autoregressive Conditional Hazard (ACH) model |
| Results | The standard indicators, as well as contagion and past exchange rate and reserves behavior, increase vulnerability. In and out of sample, 7 of 8 crises are called, but 1 of 3 tranquil periods are false alarms. | Finds significant differences between the pre- and post-liberalization regressions; chance of a currency crisis in response to overvaluation, short-term loans is greater post-liberalization | The ACH model is found to fit better than a probit model, based on a comparison of log likelihoods. None of the indicators is significant, save for the duration/contagion measure. |

REFERENCES

- Abiad, Abdul, 2002, "Early Warning Systems for Currency Crises: A Markov-Switching Approach with Application to Southeast Asia" (Ph.D. Dissertation; Philadelphia: University of Pennsylvania).
- Apoteker, Thierry, 2000, "TAC Risk Monitor: Method and Results," unpublished; available on the web at www.tac-financial.com/pdf/rmonitor_men.pdf.
- and Sylvain Barthelemy, 2001, "Genetic Algorithms and Financial Crises in Emerging Markets," unpublished; available on the web at www.tac-financial.com/publ/affi-namur-2001.pdf.
- Aziz, Jahangir, Francesco Caramazza, and Ranil Salgado, 2000, "Currency Crises: In Search of Common Elements," IMF Working Paper 00/67 (Washington: International Monetary Fund).
- Berg, Andrew, Eduardo Borensztein, and Catherine Pattillo, 2003, "Assessing Early Warning Systems: How Have They Worked in Practice?" IMF Working Paper (Washington: International Monetary Fund), forthcoming.
- Berg, Andrew, and Catherine Pattillo, 1999, "Predicting Currency Crises: The Indicators Approach and an Alternative," *Journal of International Money and Finance*, Vol. 18, (August), pp. 561–86.
- Blejer, Mario I., and Liliana Schumacher, 1998, "Central Bank Vulnerability and the Credibility of Commitments: A Value-at-Risk Approach to Currency Crises," IMF Working Paper 98/65 (Washington: International Monetary Fund).
- Bruggemann, Axel, and Thomas Linne, 2000, "Are the Central and Eastern European Transition Countries Still Vulnerable to a Financial Crisis?: Results from the Signals Approach," Bank of Finland Institute for Economies in Transition Discussion Paper (Helsinki).
- Burkart, Olivier, and Virginie Coudert, 2000, "Leading Indicators of Currency Crises in Emerging Economies," *Notes D'Etudes et de Recherche*, Banque de France (March).
- Bussiere, Matthieu, and Christian Mulder, 1999a, "Political Instability and Economic Vulnerability," IMF Working Paper 99/46 (Washington: International Monetary Fund).
- 1999b, "External Vulnerability in Emerging Market Economies: How High Liquidity Can Offset Weak Fundamentals and the Effects of Contagion," IMF Working Paper 99/88 (Washington: International Monetary Fund).

- Caramazza, Francesco, Luca Ricci, and Ranil Salgado, 2000, "Trade and Financial Contagion in Currency Crises," IMF Working Paper 00/55 (Washington: International Monetary Fund).
- Cecchetti, S.G., P. Lam, and N.C. Mark, 1990, "Mean Reversion in Equilibrium Asset Prices," *American Economic Review*, 80 (3), pp. 398–418.
- Cerra, Valerie, and Sweta Chaman Saxena, 2002, "Contagion, Monsoons, and Domestic Turmoil in Indonesia's Currency Crisis," *Review of International Economics* 10(1), pp. 36–44 (February).
- Collins, Susan M., 2001, "A Model of the Timing of Currency Crises." Georgetown University, unpublished manuscript, August.
- Davies, R. B., 1977, "Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64(2), pp. 247–54.
- , "Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74(1), pp. 33–43.
- Diebold, F. X., G. C. Weinbach, and J. H. Lee, 1994, "Regime Switching with Time-Varying Transition Probabilities," in *Nonstationary Time Series Analysis and Cointegration*, ed. by C. P. Hargreaves; Oxford University Press.
- Edison, Hali J., 2000, "Do Indicators of Financial Crises Work? An Evaluation of an Early Warning System." Federal Reserve Board of Governors, International Finance Discussion Papers; No. 675, pp. 1–74, July.
- Eichengreen, Barry, Andrew Rose, and Charles Wyplosz, 1996, "Contagious Currency Crises: First Tests," *Scandinavian Journal of Economics*, 98(4), pp. 463–84
- Eliasson, Ann-Charlotte, and Christof Kreuter, 2001, "On Currency Crisis Models: A Continuous Crisis Definition." Deutsche Bank Research Quantitative Analysis unpublished working paper; available online at www.economia.uniroma2.it/ceis/conferenze_convegni/banking2001/papers/mercoledi/Eliasson-Kreuter.pdf
- Engel, Charles, and James D. Hamilton, 1990, "Long Swings in the Dollar: Are They in the Data, and do the Markets Know it?" *American Economic Review*, 80(4), pp. 689–713.
- Esquivel, Gerardo, and Felipe Larrain, 2000, "Explaining Currency Crises," *El Trimestre Económico* 67(2), pp. 191–237 (April–June).

- Filardo, Andrew J., 1994, "Business Cycle Phases and Their Transitional Dynamics," *Journal of Business and Economic Statistics*, pp. 299–308.
- , and Stephen F. Gordon, 1993, "Business Cycle Durations." Federal Reserve Bank of Kansas Research Working Paper No. 28, pp. 93–11, October.
- Flood, Robert, and Nancy Marion, 1999, "Perspectives on the Recent Currency Crisis Literature," *International Journal of Finance and Economics*, Vol. 4, No. 1, pp. 1-26 (January)
- Frankel, Jeffrey, and Andrew Rose, 1996. "Currency Crashes in Emerging Markets: An Empirical Treatment," *Journal of International Economics*, Vol. 41, pp. 351–66 (November)
- Fratzscher, Marcel, 1999, "What Causes Currency Crises: Sunspots, Contagion or Fundamentals?" European University Institute Department. of Economics. EIU Working Papers 99/39, p. 45.
- Ghosh, Swati, and Atish Ghosh, 2002. "Structural Vulnerabilities and Currency Crises," International Monetary Fund Working Paper: WP/02/9, January.
- Goldfeld, S. M., and R. E. Quandt, 1973. "A Markov Model for Switching Regressions," *Journal of Econometrics*, 1, pp. 3–16.
- Goldstein, Morris, Graciela L. Kaminsky, and Carmen M. Reinhart, 2000, *Assessing Financial Vulnerability: An Early Warning System for Emerging Markets*. (Washington: Institute for International Economics).
- Grier, Kevin B., and Robin M. Grier, 2001, "Exchange Rate Regimes and the Cross-Country Distribution of the 1997 Financial Crisis," *Economic Inquiry* 39(1), pp. 139–48 (January).
- Hamilton, James D., 1988, "Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control*, 12, pp. 385–423.
- , 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, pp. 357–84.
- , 1990, "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics*, 45, pp. 39–70.
- , 1994, *Time Series Analysis*. Princeton University Press.

- , 1996, "Specification Testing in Markov-switching Time Series Models," *Journal of Econometrics*, 70, pp. 127–57.
- , and Oscar Jorda, 2002, "A Model of the Federal Funds Rate Target," *Journal of Political Economy* 110(5), pp. 1135–67 (October).
- Hansen, Bruce E., 1992, "The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, pp. 195–98.
- Hansen, Bruce E., 1996, "Inference when a Nuisance Parameter is Not Identified Under the Null Hypothesis," *Econometrica*, 64, pp. 413–30.
- Hawkins, John, and Marc Klau, 2000, "Measuring Potential Vulnerabilities in Emerging Market Economies." BIS Working Paper No. 91, October.
- Herrera, Santiago, and Conrado Garcia, 1999, "User's Guide to an Early Warning System for Macroeconomic Vulnerability in Latin American Countries." World Bank Working Paper 2233, November.
- Jeanne, Olivier, and Paul Masson, 2000, "Currency Crises, Sunspots and Markov-Switching Regimes." *Journal of International Economics* 50(2), pp. 327–50 (April).
- Kamin, Steven B., John W. Schindler, and Shawna L. Samuel, 2001, "The Contributions of Domestic and External Factors to Emerging Market Devaluation Crises: An Early Warning Systems Approach." Board of Governors of the Federal Reserve System, International Finance Discussion Paper No. 711 (September).
- Kaminsky, Graciela, L., Saúl Lizondo, and Carmen Reinhart, 1997, "Leading Indicators of Currency Crises," *IMF Staff Papers*, 45(1), pp. 1–48 (March).
- Kaminsky, Graciela, L., and Carmen Reinhart, 1996, "The Twin Crises: Causes of Banking and Balance-of-Payments Crises," *American Economic Review* 89(3) pp. 473–500 (June).
- Kaufmann, Daniel, Gil Mehrez, and Sergio Schmukler, 1999, "Predicting Currency Fluctuations and Crises: Do Resident Firms Have an Informational Advantage?" World Bank Working Paper No. 2259, December.
- Kim, Chang-Jin, and Charles R. Nelson, 1999, *State-space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge, Massachusetts: MIT Press.
- Krkoska, Libor, 2001, "Assessing Macroeconomic Vulnerability in Central Europe," *Post-Communist Economies* 13(1), pp. 41–55 (March).

- Kumar, Manmohan, Uma Moorthy, and William Perraudin, 2002, "Predicting Emerging Market Currency Crashes," IMF Working Paper 02/07 (Washington: International Monetary Fund).
- Kwack, Sung Yeung, 2000, "An Empirical Analysis of the Factors Determining the Financial Crisis in Asia," *Journal of Asian Economics* 11(2), pp. 195–206 (Summer).
- Lee, Joon-Haeng, 1991, "Non-Stationary Markov-Switching Models of Exchange Rates: The Pound-Dollar Exchange Rate." Ph.D. Dissertation, University of Pennsylvania.
- Mariano, Roberto S., and Fang-Xiong Gong, 1998, "Testing Under Non-standard Conditions in the Frequency Domain: With Applications to Markov Regime Switching Models of Exchange Rates and the Federal Funds Rate," Federal Reserve Bank of New York Staff Papers.
- Martinez-Peria, Maria Soledad, 2002, "A Regime Switching Approach to Studying Speculative Attacks: A Focus on EMS Crises," *Empirical Economics* 27(2), pp 299–334.
- Mulder, Christian, Roberto Perrelli, and Manuel Rocha, 2002, "The Role of Corporate, Legal and Macroeconomic Balance Sheet Indicators in Crisis Detection and Prevention," International Monetary Fund Working Paper: WP/02/59, March.
- Nag, Ashok, and Amit Mitra, 1999, "Neural Networks and Early Warning Indicators of Currency Crisis," *Reserve Bank of India Occasional Papers* 20(2), pp. 183–222 (Monsoon).
- Nitithanprapas, Ekniti, and Thomas D. Willett, 2000, "A Currency Crisis Model that Works: A Payments Disequilibrium Approach." Claremont Colleges Working Papers in Economics No. 2000–25.
- Osband, Kent, and Caroline Van Rijckeghem, 1998, "Vulnerability to Currency Crises," *IMF Staff Papers* 47(2), pp. 238–58.
- Quandt, R. E., 1958, "The Estimation of Parameters of Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, 55, pp. 873–80.
- Sachs, Jeffrey, Aaron Tornell, and Andrés Velasco, 1996, "Financial Crises in Emerging Markets: The Lessons from 1995," *Brookings Papers on Economic Activity: 1*. Brookings Institution, pp. 147–215.
- Tornell, Aaron, 1999, "Common Fundamentals in the Tequila and Asian Crises," National Bureau of Economic Research Working Paper: No. 7139, May.

- Vlaar, Peter J.G., 2000, "Currency Crises Models for Emerging Markets," De Nederlandsche Bank Staff Report No. 45 (November).
- Weller, Christian E., 2001, "Financial Crises after Financial Liberalisation: Exceptional Circumstances or Structural Weakness?" *Journal of Development Studies* 38(1), pp. 98–127 (October).
- Zhang, Zhiwei, 2001, "Speculative Attacks in the Asian Crisis," (unpublished, Working Paper, UC-San Diego, May).