THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 2: 287-307

Monitoring Banking Sector Fragility: A Multivariate Logit Approach

Aslı Demirgüç-Kunt and Enrica Detragiache

This article explores how a multivariate logit model of the probability of a banking crisis can be used to monitor banking sector fragility. The proposed approach relies on readily available data, and the fragility assessment has a clear interpretation based on in-sample statistics. The model has better in-sample performance than currently available alternatives, and the monitoring system can be tailored to fit the preferences of decisionmakers regarding type I and type II errors. The framework can be useful as a preliminary screen to economize on precautionary costs.

The past two decades have seen a proliferation of systemic banking crises, as documented by Lindgren, García, and Saal (1996) and Caprio and Klingebiel (1996), among other comprehensive studies. The spread of banking sector problems and the difficulty of anticipating their outbreak have highlighted the need to improve monitoring capabilities at both the national and supranational levels and raised the issue of using statistical studies of past banking crises to develop a set of indicators of the likelihood of future problems.

In our previous work we developed an empirical model of the determinants of systemic banking crises for a large panel of countries (Demirgüç-Kunt and Detragiache 1998, 1999). That research revealed a group of variables, including macroeconomic variables, characteristics of the banking sector, and structural characteristics of the country, that are robustly correlated with the emergence of banking sector crises. In this article we explore how we can use the information contained in that empirical relationship to monitor banking sector fragility.[1]

The basic idea is to estimate a specification of the multivariate logit model used in our previous work that relies mainly on explanatory variables whose future values are routinely forecasted by professional forecasters, the International Monetary Fund (IMF), or the World Bank. We then compute the probabili-

1. Other studies using limited dependent variable econometric models to estimate the probabilities of banking crises are Eichengreen and Rose (1998) and Hardy and Pazarbasioglu (1998). These studies do not address issues of forecasting.

Aslı Demirgüç-Kunt is lead economist with the Development Research Group at the World Bank, and Enrica Detragiache is senior economist with the Research Department at the International Monetary Fund. Their e-mail addresses are *ademirguckunt@worldbank.org* and *edetragiache@imf.org*. The authors wish to thank Anqing Shi for capable research assistance.

ties of out-of-sample banking crises using the estimated coefficients and forecasted values of the explanatory variables. Along with the results of in-sample estimations, we use these forecasted probabilities to make a quantitative assessment of fragility.

We examine two monitoring frameworks. In the first the monitor wants to know whether the forecasted probabilities are high enough to trigger a response. Taking no action when a crisis is nearing is costly, but so is taking action when a crisis is not impending. The decisionmaker chooses a probability threshold that minimizes a loss function reflecting both types of cost. In the second framework the monitor is simply interested in rating the fragility of the banking system. Depending on the rating, several courses of action may follow, but these are not explicitly modeled. In this framework it is desirable for a rating to have a clear interpretation in terms of the probability of a crisis, so that different ratings can be compared. We examine one such example.

To illustrate the monitoring procedures developed in the first part of the article, we then conduct a limited out-of-sample forecasting exercise in the second part. We construct forecasted probabilities for the six banking crises that occurred in 1996–97, namely the Jamaican crisis in 1996 and the five East Asian crises in 1997.

I. The Literature

An extensive literature reviews banking crises around the world, examining the developments leading up to the crises as well as policy responses. This body of work does not directly identify leading indicators of banking sector problems, pointing instead to a number of variables that display "anomalous" behavior in the periods preceding the crises. For instance, Gavin and Hausman (1996) and Sachs, Tornell, and Velasco (1996) suggest that credit growth be used as an indicator of impending troubles, as crises tend to be preceded by lending booms. Mishkin (1996) highlights equity price declines, while Calvo (1996), in his analysis of Mexico's 1995 crisis, suggests that monitoring the ratio of broad money to foreign exchange reserves may be useful in evaluating the banking sector's vulnerability to a currency crisis.

Honohan (1997) evaluates alternative indicators more systematically. Using a sample of 18 countries that experienced banking crises and 6 that did not, he divides the crisis countries into three groups (of equal size) according to the type of crisis (macroeconomic, microeconomic, or related to the behavior of the government). He then compares the average values of seven indicators for the crisis countries with the averages for the control group. This exercise shows that banking crises arising from macroeconomic problems are associated with high loan-to-deposit ratios, high foreign borrowing-to-deposit ratios, and high growth rates of credit. Similarly, crises stemming from government interventions are associated with high levels of government borrowing and central bank lending to the banking system. However, banking crises originating from microeconomic pres-

sures do not appear to be associated with abnormal behavior on the part of the indicators examined in the study.

Rojas-Suárez (1998) proposes an approach based on bank-level indicators, similar in spirit to the CAMEL system used by U.S. regulators to identify problem banks. She argues that in emerging markets (particularly those in Latin America) CAMEL indicators are not good signals of bank strength and that more information can be obtained by monitoring the deposit interest rate, the spread between the lending and deposit rates, the growth rate of credit, and the growth rate of interbank debt. Because these variables are measured against banking system averages, however, this approach appears more adequate for identifying weaknesses specific to individual banks than for identifying systemic fragility. The approach also requires bank-level information, which often is not readily available in developing countries.

To date, Kaminsky and Reinhart (1999) have made the most comprehensive effort to develop a set of early warning indicators for banking crises (and currency crises). The methodology is refined in Kaminsky (1998). These studies examine the behavior of 15 macroeconomic indicators for a sample of 20 countries that experienced banking crises during 1970–95.[2] The authors compare the behavior of each indicator in the 24 months prior to the crisis with the behavior during tranquil times. A variable is deemed to signal a crisis if it crosses a particular threshold at any time. If that signal is followed by a crisis within the next 24 months, then it is considered correct; otherwise it is considered noise. The threshold for each variable is chosen to minimize the in-sample noise-to-signal ratio. The authors then compare the performance of different indicators based on the associated type I and type II errors, the noise-to-signal ratio, and the probability of a crisis occurring conditional on a signal being issued.[3]

The indicator with the lowest noise-to-signal ratio and the highest probability of crisis conditional on the signal is the real exchange rate, followed by equity prices and the money multiplier. These three indicators, however, have a high incidence of type I errors, as they fail to issue a signal in 73–79 percent of the observations in the 24 months preceding a crisis. The incidence of type II errors, in contrast, is much lower, ranging from 8 to 9 percent. The variable with the lowest type I error is the real interest rate, which issues a signal in 30 percent of the observations preceding a crisis. The high incidence of type I errors relative to type II errors may not be a desirable feature of a warning system if the costs of raising a false alarm are small relative to the costs of failing to anticipate a crisis.

Since, presumably, the likelihood of a crisis is greater when several indicators signal simultaneously, Kaminsky (1998) develops composite indexes. These include the number of indicators that cross the threshold at any given time or a weighted variant of that index in which each indicator is weighted by its signal-to-noise ratio so that more informative indicators receive more weight. The best

2. For a study of early warning indicators of currency crises, see also IMF (1998).

3. The authors use an adjusted version of the noise-to-signal ratio, computed as the ratio of the probability of a type II error to 1 minus the probability of a type I error.

composite indicator outperforms the real exchange rate in predicting crises in the sample, but it is worse at predicting observations of no crisis.[4]

The approach we develop here will allow policymakers to choose a warning system that reflects the relative cost of type I and type II errors, and it will offer a natural way of measuring the combined effect of various economic forces on banking sector vulnerability. By making better use of all available information, the system will produce lower overall in-sample forecasting errors than would individual indicators. We also examine a problem not addressed by Kaminsky and Reinhart (1999), that of a monitor who wishes to use information contained in the statistical analysis of past crises not just to anticipate a crisis but also to make a more nuanced assessment of banking sector fragility.

II. Estimating the Probabilities of In-Sample Banking Crises in a Multivariate Logit Framework

The starting point of our analysis is an econometric model of the probability of a systemic banking crisis. In Demirgüç-Kunt and Detragiache (1998, 1999) we estimate alternative specifications of a logit regression for a large sample of developing and industrial countries, including countries that experienced banking crises and those that did not. Details on sample selection, the construction of the banking crisis variable, and the choice of explanatory variables can be found there.

To form the basis of an easy-to-use monitoring system, we estimate a specification of our empirical model that includes only variables that are available from the IMF's *International Financial Statistics* or other publicly available databases and that are routinely forecasted by the IMF in its biannual *World Economic Outlook* or by professional forecasters. As it turns out, this is not the specification that best fits the data. We estimate the regression using a panel of 766 observations for 65 countries during 1980–95.[5] In this panel we identify 36 systemic banking crises, so that crisis observations make up 4.7 percent of the sample (table 1). The set of explanatory variables capturing macroeconomic conditions includes the growth rate of real gross domestic product (GDP), the change in the terms of trade, the rate of depreciation of the exchange rate (relative to the U.S. dollar), the rate of inflation, and the fiscal surplus as a share of GDP. The explanatory variables capturing characteristics of the financial sector are the ratio of broad money to foreign exchange reserves and the growth rate of bank credit lagged two periods. Finally, GDP per capita proxies for structural characteristics of the economy.

4. Kaminsky (1998) finds that the probability of a crisis computed by taking into account the number of indicators signaling a crisis increased substantially before the 1997 crises in the Philippines, Malaysia, and Thailand, but not in Indonesia. The Republic of Korea was not part of the sample.

5. Because of missing data or breaks in the series, part of the sample period may be excluded for some countries. Years in which banking crises are ongoing also are excluded from the sample.

Table 1. *Banking Crises and Estimated Probabilities of Crises*

Country	*Crisis year*	*Estimated probability*
Chile	1981	0.231
Colombia	1982	0.066
Ecuador	1995	0.439
El Salvador	1989	0.055
Finland	1991	0.066
Guyana	1993	0.007
India	1991	0.069
Indonesia	1992	0.107
Israel	1983	0.999
Italy	1990	0.015
Japan	1992	0.037
Jordan	1989	0.334
Kenya	1993	0.361
Malaysia	1985	0.067
Mali	1987	0.035
Mexico	1982	0.527
Mexico	1994	0.099
Nepal	1988	0.018
Nigeria	1991	0.011
Norway	1987	0.036
Panama	1988	0.539
Papua New Guinea	1989	0.121
Peru	1983	0.244
Philippines	1981	0.035
Portugal	1986	0.064
South Africa	1985	0.196
Sri Lanka	1989	0.036
Swaziland	1995	0.633
Sweden	1990	0.036
Tanzania	1988	0.035
Thailand	1983	0.027
Turkey	1991	0.158
Turkey	1994	0.482
United States	1980	0.238
Uruguay	1981	0.329
Venezuela	1993	0.494

Source: Authors' calculations.

The estimated coefficients of the logit regression reveal that low GDP growth, a high real interest rate, high inflation, strong growth of bank credit in the past, and a high ratio of broad money to reserves are all associated with a high probability of a banking crisis (table 2). Exchange rate depreciation, the terms of trade, the fiscal surplus, and GDP per capita are not significant. The estimated probability of a crisis for the 36 episodes included in the sample ranges from a low of 1.1 percent for Nigeria to a high of 99.9 percent for Israel (see table 1). About 70 percent of the episodes have an estimated probability of 4 percent or more, while only 17 percent have an estimated probability of more than 50 percent.

Table 2. *Logit Regression of the Probability of a Banking Crisis*

Explanatory variable	*Estimated coefficient*
GDP Growth	–0.172*
	(0.034)
Change in terms of trade	–0.021
	(0.018)
Depreciation	0.007
	(0.006)
Real interest rate	0.065*
	(0.016)
Inflation	0.020**
	(0.010)
Ratio of fiscal surplus to GDP	0.066
	(0.036)
Ratio of M2 to reserves	0.013*
	(0.005)
Credit growth $_{t-2}$	0.015**
	(0.008)
GDP per capita	–0.039
	(0.033)
Number of crises	36
Number of observations	766
Model χ^2	61.46*
AIC[a]	249

*Significant at the 1 percent level.
**Significant at the 5 percent level.
Note: Standard errors are in parentheses.
a. Akaike's Information Criterion.
Source: Authors' calculations.

Sources of Fragility: The 1994 Mexican Crisis

One of the advantages of the multivariate logit model is that we can easily identify the sources of fragility by calculating the contribution of each explanatory variable to a change in the estimated probability of a crisis. As an illustration, we analyze the factors that contributed to the sharp increase in the estimated probability of a crisis in Mexico in 1993, just before the actual crisis occurred in 1994 (table 3).

In 1993 high past credit growth, high real interest rates, and high inflation were the main factors underlying the high probability of a crisis in Mexico. Because the logit is nonlinear, the sum of the contribution of each variable to the change in probability does not always add up to the total change (see the last column of table 3). Looking at macroeconomic factors, we see that Mexico had a negative growth shock that significantly raised the probability of a crisis. Real interest rates also rose significantly, and there was a minor terms-of-trade shock. At the same time, appreciation of the exchange rate, lower inflation, and a lower budget surplus offset some of this increase.

Financial sector variables played a less important role in explaining the overall increase in probability, slightly offsetting the impact of the macroeconomic fac-

Table 3. *Decomposition of the Estimated Probability of a Banking Crisis, Mexico, 1992–93*

Explanatory variable	*Percentage change in variable, 1992–93*	*Weight in 1993*[a]	*Weight in 1992*[a]	*Change in weight, 1992–93*	*Contribution to percentage change in probability*
GDP growth	–125	0.154	–0.624	0.778	105
Change in terms of trade	–16	–0.034	–0.041	0.007	1
Depreciation	–119	–0.002	0.010	–0.012	–1
Real interest rate	386	0.327	0.067	0.259	28
Inflation	–31	0.202	0.295	–0.093	–8
Ratio of fiscal surplus to GDP	–79	0.022	0.102	–0.080	–7
Ratio of M2 to reserves	–16	0.057	0.068	–0.011	–1
Credit growth$_{t-2}$	–4	0.498	0.517	–0.019	–2
GDP per capita	–1	–0.070	–0.070	0.000	0
Estimated probability of a crisis					
1992	0.054				
1993	0.116				

a. Weights are obtained by multiplying the estimated regression coefficient of each variable by the value of the variable. A negative weight indicates that the variable reduced the estimated probability of a crisis.

Source: Authors' calculations.

tors. The vulnerability of the financial system to capital outflows—measured by M2 divided by the reserves ratio—fell slightly, leading to a 1 percent decrease in the probability of a crisis. Credit growth slowed, reducing the probability by 2 percent. Finally, GDP per capita—which we use as a proxy of institutional development—did not change significantly in this period. Thus decomposing the probability of a crisis helps us to understand which factors played a role in bringing about the crisis, at least according to the empirical model.

Out-of-Sample Probability Forecasts

Because the purpose of monitoring is to assess future fragility, the next step is to forecast the probability of a banking crisis. Let $\boldsymbol{\beta}$ be a $1 \times N$ vector containing the N estimated coefficients of the logit regression reported in table 1, and let $\mathbf{z}_{it}$ be an $N \times 1$ vector of out-of-sample values of the explanatory variables for country i at date t. These values can be true forecasts, estimates of past values, data for countries or time periods not included in the sample, or ranges of values to construct alternative scenarios. Then the out-of-sample probability of a banking crisis for country i at date t is

$$p_{it} = \frac{\exp[\boldsymbol{\beta}\, \mathbf{z}_{it}]}{1 + \exp[\boldsymbol{\beta}\, \mathbf{z}_{it}]}. \tag{1}$$

Once we compute out-of-sample probabilities, the question arises of how to interpret them. Is a 10 percent probability of a crisis high or low? Should a

policymaker take preventive actions when faced with such a probability? Should a surveillance agency issue a warning? In the next section we address such questions.

III. Building an Early Warning System Using Estimated Crisis Probabilities

The first monitoring framework that we consider is one in which the decisionmaker must decide whether the forecasted probability is large enough to issue a warning. This is the framework implicit in Kaminsky and Reinhart (1999). Issuing a warning will lead to some sort of preventive action. For instance, the decisionmaker may invest in gathering further information, such as acquiring bank-level balance sheet data or holding discussions with senior bank managers, bank supervisory agencies, or other market participants. Alternatively, the decisionmaker may use the monitoring system to decide whether to take preventive policy measures, such as tightening prudential capital or liquidity requirements for banks or reducing interest rates to ease pressures on bank balance sheets. For a warning system to be useful, preventive measures must substantially reduce the costs of a crisis. We assume that this is the case. Also a useful warning system should minimize false alarms, since preventive measures are usually costly. Tighter prudential requirements may cause banks to cut credit, perhaps leading to a credit crunch; looser monetary policy may lead to higher inflation.

The choice of the threshold for issuing a warning will generally depend on three factors. The first is the probability of type I and type II errors associated with the threshold, which, assuming that the sample of past crises is representative of future crises, can be assessed on the basis of the in-sample frequency of the two errors. Clearly, the higher is the threshold that forecasted probabilities must cross before a warning is issued, the higher will be the probability of a type I error and the lower will be the probability of a type II error (and vice versa).

The second parameter on which the choice of the threshold depends is the unconditional probability of a banking crisis, which can also be assessed based on the in-sample frequency of crisis observations. If crises tend to be rare events, then the overall likelihood of making a type I error is relatively small (and vice versa). Finally, the third factor is the cost to the decisionmaker of taking preventive actions relative to the cost of failing to anticipate a banking crisis. In general, these costs to the decisionmaker are themselves forecasts of the true costs, and making a good decision requires having good forecasts. A policymaker who tends to underestimate the cost of a crisis or to overestimate the cost of taking preventive actions will be too conservative in choosing a warning threshold (and vice versa).[6]

A Loss Function for the Decisionmaker

Based on these considerations, we can develop a more formal analysis of the decision process behind the choice of a warning system. Let T be the threshold

6. For estimates of the fiscal costs of recent banking crises, see Caprio and Klingebiel (1996).

chosen by the decisionmaker, so that if the forecasted probability of a crisis for country i at time t exceeds T, the system will issue a warning. Let $p(T)$ denote the probability that the system will issue a warning, and let $e(T)$ be the joint probability that a crisis will occur and the system will not issue a warning. Further, let c_1 be the cost of taking preventive actions as a result of having received a warning, and let c_2 be the additional cost of a banking crisis if it is not anticipated (if anticipating a crisis can prevent it altogether, then c_2 is the entire cost of the crisis). Presumably, c_1 is substantially smaller than c_2 if further information gathering will be useful and if the knowledge that a crisis is impending will allow policymakers to take effective preventive measures. Then we can define a simple linear expected loss function for the decisionmaker as

$$L(T) \equiv p(T)c_1 + e(T)c_2. \tag{2}$$

Let $a(T)$ be the type I error associated with threshold T (the probability of not receiving any warning conditional on a crisis occurring), and let $b(T)$ be the probability of a type II error (the probability of receiving a warning conditional on no crisis taking place). Also let w denote the (unconditional) probability of a crisis. Then we can rewrite the loss function of the decisionmaker as

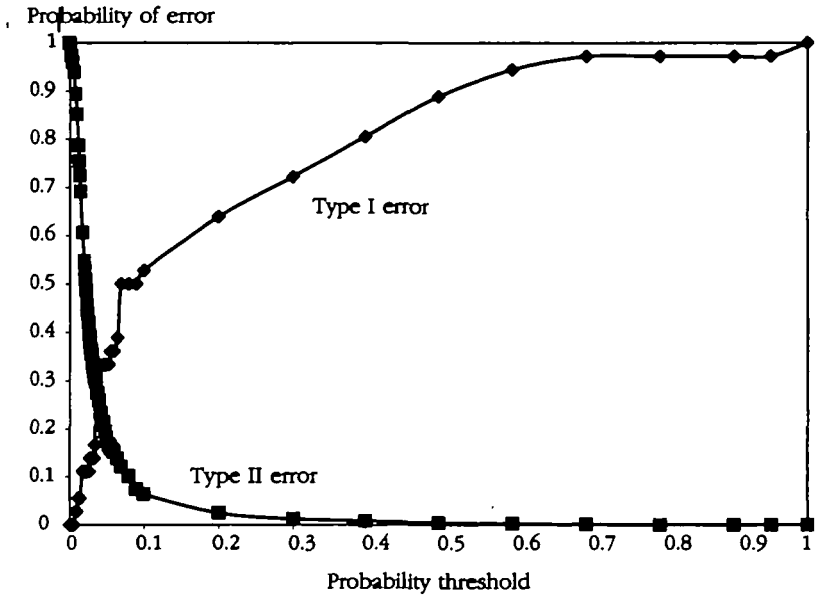
$$\begin{aligned} L(T) &= c_1[(1-a(T))w + b(T)(1-w)] + c_2 a(T)w \\ &= wc_1\left[1 + \left(\frac{c_2 - c_1}{c_1}\right)a(T) + b(T)\left(\frac{1-w}{w}\right)\right]. \end{aligned} \tag{3}$$

The second part of the equality shows that the higher is the cost of missing a crisis relative to the cost of taking preventive action (the larger is c_2 relative to c_1), the more concerned will the decisionmaker be about a type I error relative to a type II error (and vice versa). Also the higher is the unconditional probability of a banking crisis (measured by the parameter w), the more weight will the decisionmaker place on type II errors, as the frequency of false alarms is greater when crises tend to be rare events.[7]

Using in-sample frequencies as estimates of the true parameters, w should equal the frequency of banking crises in the sample, namely 0.047 (see table 1). We can obtain the functions $a(T)$ and $b(T)$, which trace how error probabilities change with the threshold for issuing warnings, from the in-sample estimations as follows. Given a threshold of, say, $T = 0.05$, we can derive $a(0.05)$, that is, the associated probability of a type I error, as the percentage of banking crises in the sample with an estimated probability below 0.05. Similarly, $b(0.05)$, the probability of issuing a warning when no crisis occurs, is the percentage of observations in which no crisis occurs when the estimated probability of a crisis is above 0.05. For $T \in [0, 1]$, $a(T)$ is increasing, since the probability of not issuing a

7. A risk-averse decisionmaker would place greater weight on minimizing type I errors than on minimizing type II errors, since type I errors are more costly. We are indebted to a referee for suggesting this point.

Figure 1. *Crisis Threshold and In-Sample Classification Accuracy*



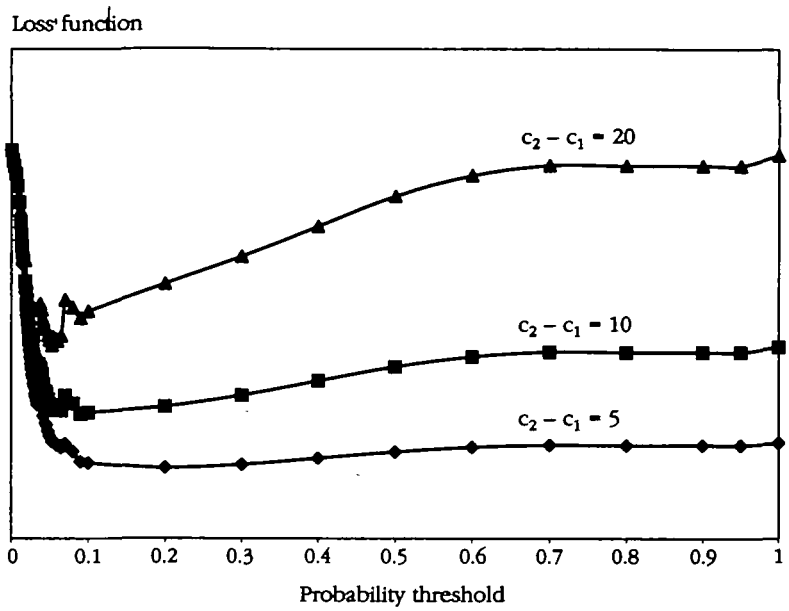
Source: Authors' calculations.

warning when a crisis occurs increases as the threshold rises, while $b(T)$ is decreasing (figure 1). The two functions cross at $T = 0.036$, at which the probability of either type of error is about 30 percent.

Figure 1 also shows that probabilities estimated through our multivariate logit framework can provide a more accurate basis for an early warning system than the indicators developed by Kaminsky and Reinhart (1999). The indicator associated with the lowest type I error in the Kaminsky-Reinhart framework is the real interest rate, with a type I error of 70 percent and a type II error of 19 percent. In our model a threshold for a type I error of 72 percent comes at the cost of a type II error of only 1.2 percent. Similarly, the best indicator of banking crises according to Kaminsky and Reinhart is the real exchange rate, with a type I error of 73 percent and a type II error of 8 percent (resulting in an adjusted noise-to-signal ratio of 0.30). With our model we can obtain a type II error of 7.4 percent by choosing a probability threshold of 0.09, which is associated with a type I error of only 53 percent. The adjusted noise-to-signal ratio is 0.25. The better performance of the multivariate logit model likely stems from the fact that it combines into one number (the estimated probability of a crisis) all of the information provided by the economic variables monitored.[8]

8. The logit parameters are estimated using maximum likelihood, and the likelihood function does not take into account the different costs of type I and type II errors. One way to improve the warning system could be to choose parameters that minimize the decisionmaker's loss functions.

Figure 2. *Loss Functions for Varying Cost Parameters*



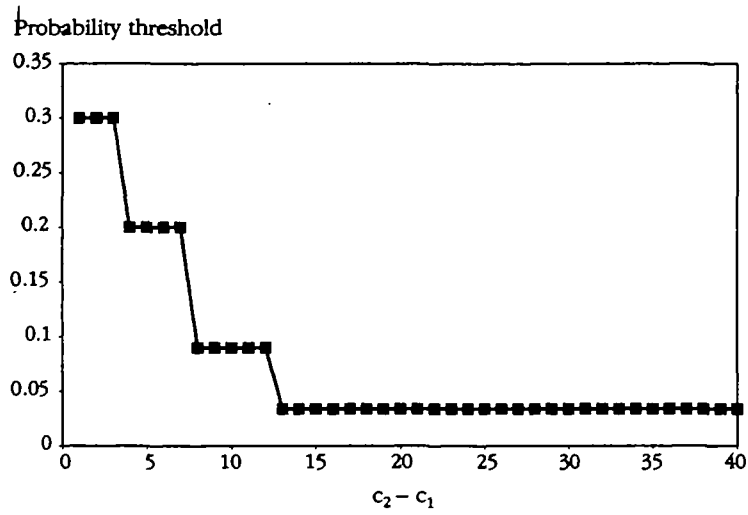
Note: The value $c_2 - c_1$ measures the cost to the decisionmaker of failing to identify a crisis relative to the cost of taking precautionary measures.
Source: Authors' calculations.

Choosing the Optimal Threshold

To illustrate, we compute loss functions for three configurations of the decisionmaker's cost parameters (figure 2). We normalize the parameter c_1 to 1 in all three scenarios and give $c_2 - c_1$ the values 20, 10, and 5. The values of the warning threshold that minimize the loss functions are, respectively, $T = 0.034$, 0.09, and 0.20. In other words, a decisionmaker whose cost of missing a crisis is, for example, 10 times the cost of taking precautionary measures will issue an alarm every time the forecasted probability of a crisis exceeds 9 percent. Thus, as expected, as the cost of missing a crisis increases relative to the cost of taking preventive action, the optimal threshold falls, resulting in a warning system with fewer type I errors and more type II errors.

For values of c_2 between 40 and 15, keeping c_1 constant at 1, the optimal probability threshold for issuing a warning is $T = 0.034$ (figure 3). With this criterion the probability of not issuing a warning when a crisis occurs is about 14 percent, while the probability of mistakenly issuing a warning is 31 percent. As c_2 falls below 15, the threshold increases to 0.09 (type I error of 50 percent and type II error of 7.4 percent) and remains there until c_2 reaches 8. At that point the threshold jumps to 0.20, as the decisionmaker becomes very concerned about false alarms. Finally, if the cost of missing a crisis is as low as two to three times

Figure 3. *Optimal Probability Thresholds for Varying Cost Parameters*



Note: The value $c_2 - c_1$ measures the cost to the decisionmaker of failing to identify a crisis relative to the cost of taking precautionary measures.
Source: Authors' calculations.

that of issuing a false warning, then the optimal threshold is 0.30, corresponding to a type I error as high as 72.2 percent and a type II error as low as 1.2 percent.

To fully appreciate the nature of the warning system, it is worth pointing out that the probability of a type I error is *not* the probability of missing a crisis. To obtain the probability of missing a crisis, we must multiply the probability of a type I error by the unconditional probability of a crisis, which in our sample is 0.047. Similarly, the probability of issuing a false warning is the size of a type II error multiplied by the frequency of noncrisis observations. With a threshold of $T = 0.09$, the probability of missing a crisis is, therefore, only 2.3 percent, since crises occur rarely. In contrast, the probability of issuing a false alarm is 7.1 percent, because observations of no crisis tend to be the majority. Thus warning systems associated with a relatively low incidence of type I errors (below 15 percent) give rise to many false alarms, in part because crises are infrequent events. If the system is used as a preliminary screen, and further information gathering can help to sort out cases in which the banking system is sufficiently sound, then the decisionmaker will accept the high incidence of type II errors.

In some cases what the model considers to be a false alarm may actually be a useful signal. To illustrate this point, we examine the false alarms generated in-sample by a threshold of 0.047. As it turns out, in 21 cases the false alarm occurred in the two years immediately preceding a crisis, suggesting that the conditions that eventually led to a full-fledged crisis were in place (and were detectable) a few years in advance. In other cases the false alarms may have corresponded to episodes of fragility that were not sufficiently severe to be classified as full-fledged crises in our empirical study. Or they may have corresponded to episodes in

which a crisis was prevented by a prompt policy response. Thus assessing the accuracy of the warning system based on the accuracy of in-sample classification may exaggerate the incidence of type II errors. However, out-of-sample predictions are subject to additional sources of error relative to in-sample predictions: the forecasted values of the explanatory variables include forecast errors, and there may be structural breaks in the relationship between banking sector fragility and the explanatory variables, making predictions based on past behavior inadequate. Also, despite the large size of our panel, the number of systemic banking crises (36) is still relatively small, so that small-sample problems may affect the estimation results. As more data become available and the size of the panel is extended, this problem should become less severe.

Comparing the Loss Function with the Noise-to-Signal Ratio

It is of interest to compare the optimal threshold derived from minimizing the loss function proposed here with the optimal threshold that would result from minimizing the (adjusted) noise-to-signal ratio, the criterion used by Kaminsky and Reinhart (1999). Define the noise-to-signal ratio as

$$NS(T) \equiv \frac{a(T)}{1-b(T)}. \tag{4}$$

Then the loss function can be rewritten as

$$L(T) = wc_1 + w(c_2 - c_1)a(T) + (1-w)c_1 NS(T)[1-a(T)], \tag{5}$$

and the first-order condition for the minimization of the loss function is

$$\frac{dL(T)}{dT} \equiv [w(c_2 - c_1) - (1-w)c_1 NS(T)]a'(T) + (1-w)c_1[1-a(T)]N'S(T) = 0. \tag{6}$$

Suppose T^{ns} is the threshold that minimizes the noise-to-signal ratio. Then, at $T = T^{ns}$, $NS'(T) = 0$, and the derivative of the loss function is

$$\frac{dL(T)}{dT}\Big|_{T=T^{ns}} \equiv [w(c_2 - c_1) - (1-w)c_1 NS(T)]a'(T). \tag{7}$$

For a convex loss function a positive (negative) sign for equation 7 means that the threshold T^{ns} is too large (small) relative to the threshold that would minimize the loss function. Accordingly, if equation 7 is positive (negative), by minimizing the noise-to-signal ratio, the decisionmaker will make too many type I (type II) errors relative to the threshold that minimizes the loss function. Since $a'(T) > 0$ (the probability of a type I error is increasing in the threshold), equation 7 has the sign of the term in square brackets. This term is more likely to be positive the larger is the cost of a type I error $(c_2 - c_1)$ relative to the cost of a type II error (c_1) and the larger is the unconditional probability of a crisis. Thus if banking crises tend to be rare, and the cost of missing a crisis is high relative to

the cost of raising a false alarm, minimizing the noise-to-signal ratio is likely to yield a choice criterion that results in too many missed crises for a decisionmaker whose preferences are captured by the linear loss function of equation 5.

IV. Constructing a System for Rating Bank Fragility

In this section we consider the problem of a monitor who must rate the fragility of a given banking system. Other agents will then use the rating to decide on a possible policy response, but the monitor is not necessarily aware of the costs and benefits of such policy actions. Another rationale for using fragility classes instead of a critical threshold as a monitoring device is that small changes in the critical threshold may lead to substantial differences in type I and type II errors, as seen in figure 1. In constructing fragility classes, the classification criterion should have a clear interpretation in terms of type I and type II errors. This has two advantages: first, agents who learn the rating can make their own cost-benefit calculations when they decide whether or not to take action, and, second, the fragility of two systems that are assigned different ratings can be compared based on a clear metric.

The starting point is once again the set of forecasted crisis probabilities obtained using the coefficients estimated in the multivariate logit regression. Clearly, a country with a forecasted probability of x should be deemed more fragile than one with an estimated probability of $y < x$. To establish fragility classes, we can partition the interval [0, 1], which is the set of possible forecasted probabilities, into a number of subintervals and assign a rating to all estimated probabilities within a given class.

There are no objective criteria for choosing one partition over another, but a number of considerations help to narrow the choices. First, because the frequency of crises in the sample is low, choosing a fine partition would give misleading results, because many classes would have no observed crises. For instance, in our sample there are no episodes with an estimated crisis probability between 4 and 5 percent, whereas there are episodes with an estimated probability between 3 and 4 percent (see figure 1). If we choose the intervals [0.04–0.05] and [0.03–0.04] as two of the classes, then it would appear that fragility decreases with the estimated probability of a crisis—an obviously misleading conclusion.

Another caveat is that the empirical distribution of the estimated probabilities is strongly skewed toward zero: only 8.5 percent of the observations have probabilities higher than 10 percent, and more than 45 percent are in the 0–2 percent range. Thus partitioning the unit interval by subsets of the same size would assign an uneven number of observations to each class, with very few observations in the highest probability intervals.

Based on these considerations, we construct a rating system with four fragility classes (table 4). We choose the upper bounds of each of the four classes so that the type I error associated with the bounds are 10, 30, 50, and 100 percent, respectively. According to this criterion, observations with forecasted probabili-

ties below 1.8 percent belong to the lowest fragility class. Observations with probabilities between 1.8 and 3.6 percent are in the second class, up to 7 percent are in the third class, and above 7 percent are in the highest class. The values of the type II error associated with the upper bound of each class are (about) 60, 30, 12, and 0 percent, respectively.

To illustrate the meaning of the fragility groupings, consider that if all observations with forecasted probabilities in classes higher than the most fragile (that is, observations with probabilities higher than 1.8 percent) were treated as crises, the likelihood of missing a crisis (given that one takes place) would be less than 10 percent. However, the probability of falsely predicting a crisis would be higher than 60 percent. Another way to put it is that 90 percent of the crisis observations in the sample have a probability higher than the probabilities in the lowest fragility class. Similarly, if one were to classify as crises only observations with forecasted probabilities in the two highest fragility classes, then the probability of missing a crisis would rise to 30 percent, and the probability of a false alarm would fall to 30 percent.

As an additional measure of the degree of fragility associated with each class, we compute the fraction of sample observations in each class that corresponds to an actual banking crisis. This measure ranges from 1.5 percent for the lowest fragility class to 16.8 percent for the highest. Thus the likelihood that an observation in the highest fragility class is a crisis is 16.8 percent; this figure may seem low, but it should be compared to the unconditional probability of a crisis: 4.7 percent (the sample frequency of crises). To put it another way, finding that the probability of a crisis falls in the highest fragility class tells the analyst that the observation is 3.5 times more likely to correspond to a crisis than the average observation. Clearly, these rating systems are just examples of many possible alternatives, and depending on the purposes of the monitor, one alternative may be preferred to another. What is important is that potential users understand the meaning of the fragility score and the criteria used in rating.

V. Applying the System to the Banking Crises of 1996–97

To gauge the performance of our monitoring mechanisms, we consider how accurately they would have predicted the six banking crises that took place in

Table 4. *A Rating System for Banking Sector Fragility*

Class	*Probability interval*	*Type I error*	*Type II error*	*Number of observations*	*Crisis per observation*
I	0.000–0.018	0.00–0.10	1.00–0.60	291	0.01
II	0.018–0.036	0.10–0.30	0.60–0.30	232	0.03
III	0.036–0.070	0.30–0.50	0.30–0.12	136	0.05
IV	0.070–1.000	0.50–1.00	0.12–0.00	107	0.17

Note: Class I is the lowest fragility class and class IV is the highest.
Source: Authors' calculations.

1996–97, that is, after the end of the sample period used in the estimation exercise above. The banking crises occurred in Jamaica in 1996 and in Indonesia, the Republic of Korea, Malaysia, the Philippines, and Thailand in 1997. Early accounts and analyses of the events surrounding the five Asian crises can be found, for instance, in IMF (1997), Radelet and Sachs (1998), and Goldstein and Hawkins (1998).

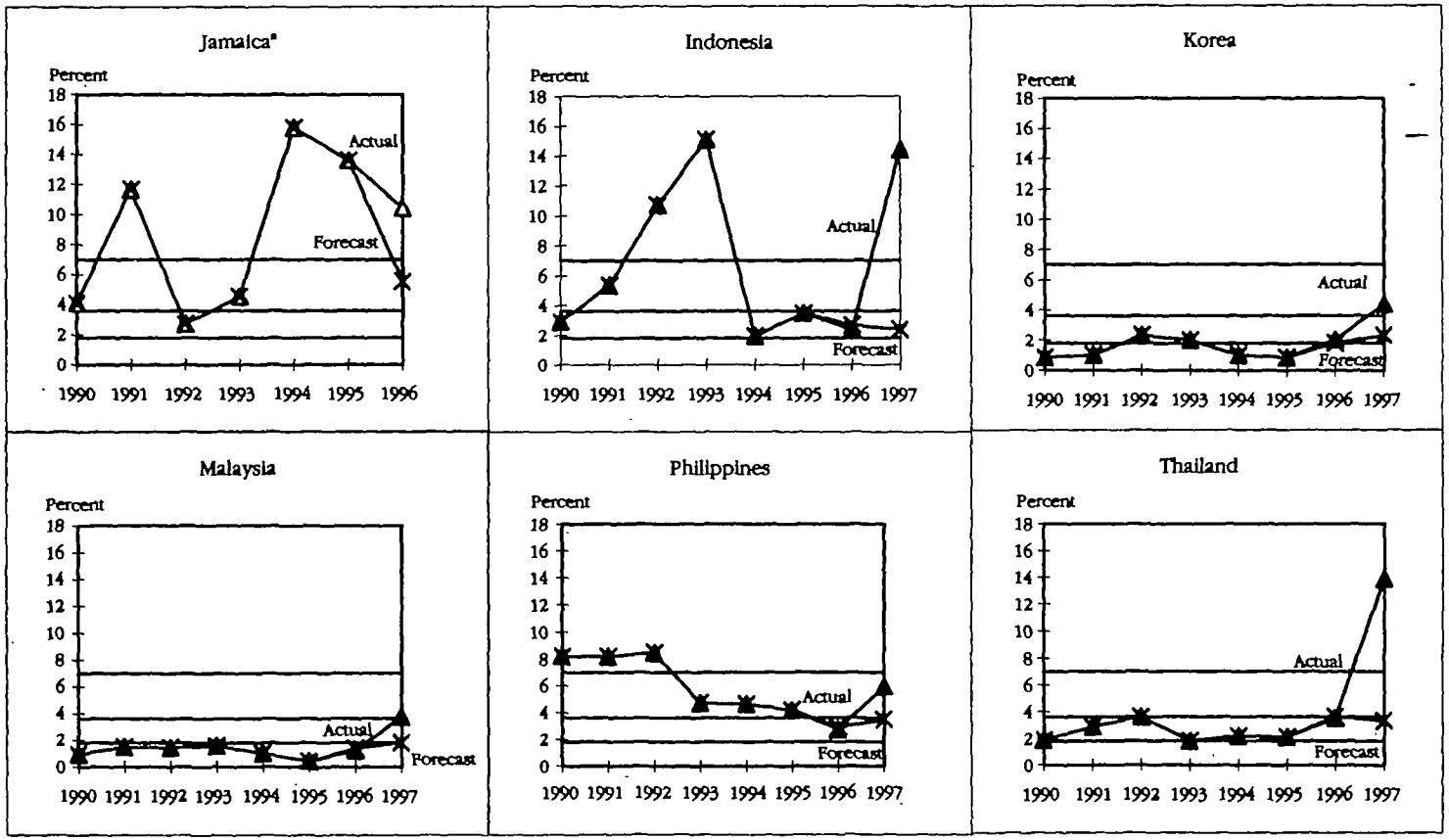
To compute the probabilities of out-of-sample banking crises for the six countries, we use two sets of values for the explanatory variables. The first set consists of actual realizations. The out-of-sample probabilities obtained in this way are not true forecasts, of course. In particular, for the five Asian countries these figures capture the large exchange rate depreciations that took place in the second half of 1997 and their immediate consequences. It is of interest to try to assess whether signs of increasing banking sector fragility would have been apparent before the depreciations took place, since they were largely unanticipated by observers. To this end, and, more generally, to assess the performance of the monitoring system when true forecasts are used, we also compute out-of-sample probabilities using forecasts of the explanatory variables as of April–May 1997. Comparing the two forecasts will reveal the extent to which errors in forecasting the explanatory variables would have clouded the fragility assessment based on our model.

We take the forecasted values of the explanatory variables, where available, from the *Financial Times' Currency Forecaster,* and from *Consensus Forecasts.* These works survey several prominent private sector forecasters and publish the means of their forecasts. For the five Asian countries the growth rate of real GDP, inflation, exchange rate depreciation, and the real interest rate are from the *Currency Forecaster*, and broad money is from *Consensus Forecasts.* The remaining values (and all of the values for Jamaica) are from the May 1997 round of the IMF's semiannual *World Economic Outlook.*[9] To compute the probabilities of out-of-sample crises using realized values of the explanatory variables, we use numbers from the *International Financial Statistics* when available and February 1998 numbers from the *World Economic Outlook* otherwise.

Based on forecasts as of April–May 1997, the estimated probabilities of crises were relatively low for the five Asian countries, whereas Jamaica was well into the highest fragility zone as early as 1995 (figure 4). This is not surprising, since all the Asian countries had very good macroeconomic performances in the years up to 1996—performances that, by and large, were expected to continue. In Jamaica the forecasted probability of a crisis was 14 percent in 1995 and 13 percent in 1996. The two main factors contributing to the increase in the probability of a crisis were high real interest rates and high inflation. Strong past

9. There are two exceptions. For the Philippines, broad money comes from the *World Economic Outlook*. For Korea, no forecast of reserves was available, so we arbitrarily assumed that reserves returned to their 1995 value in 1997.

Figure 4. *Actual and Forecasted Crisis Probabilities in Five Asian Countries and Jamaica, 1990–97*



a. Jamaican data run through 1996, the year of that country's crisis.
Source: Authors' calculations.

credit growth and a favorable fiscal position also contributed to fragility in 1995, but not in 1996.

The two most fragile Asian countries were Thailand and the Philippines, both having a forecasted crisis probability of about 3.5 percent in 1997. This probability would have placed the two countries on the border between the second and third fragility zones based on our rating system. In Thailand the main factor contributing to bank fragility both in 1996 and in 1997 was the high real interest rate; strong past credit growth was also a factor. But in contrast with Jamaica, where GDP growth was lackluster, Thailand had a large predicted GDP growth rate, which worked as an offsetting factor, keeping the overall probability of a crisis relatively low. In the Philippines the predicted probability increased more than 20 percent between 1996 and 1997, mainly because of the high growth rate of credit two years earlier. The real interest rate was lower than that in Thailand, but so was GDP growth.

Indonesia, Malaysia, and Korea all had forecasted crisis probabilities below 3 percent in 1996 and in 1997, and would have been placed in the second fragility class (actually, Malaysia would have received the lowest fragility rating in 1996). As in Thailand and the Philippines the expectation that the exchange rate would remain stable and, especially, that GDP growth would continue to be strong more than offset the prospect of fragility coming from high real interest rates (except in Korea) and strong past credit expansion. Indonesia's high rate of inflation also tended to increase bank fragility.

Not surprisingly, the picture obtained by estimating the probabilities of crises using the latest available data would have been quite different for the five Asian countries, but not for Jamaica. The estimated probabilities of crises are in the highest fragility class for Indonesia and Thailand and in the second highest for the other three Asian countries. Malaysia, with a probability of 3.7 percent, appears to have been the least fragile.[10]

Decomposing the probability tells some interesting stories. Of course, the exchange rate depreciation directly affected fragility in all five countries. However, in 1997 inflation was not much higher than forecasted, so it was not among the main factors contributing to greater banking system vulnerability. In all five countries except Korea lower-than-forecasted GDP growth was one of the main contributing factors, as was the higher-than-expected real interest rate (except in Thailand).

To summarize, an analysis of banking system fragility using the methods developed in this article would have clearly indicated an impending banking crisis in Jamaica. But although signs of fragility were present in Thailand and the Philippines, the overall image of the five Asian economies would have been fairly reassuring, as expectations of continued strong economic growth and stable exchange rates would have offset the negative impact of relatively high real interest rates and strong past credit expansion.

10. Of the five Asian countries, Malaysia is the only one without an IMF program.

VI. Conclusions

Econometric analysis of systemic banking crises is a relatively new field of study, and the development and evaluation of monitoring and forecasting tools based on the results of such analyses are at an embryonic stage at best. The purpose of this article has been not so much to propose one or more "ready-to-use" procedures for decisionmakers, but rather to highlight which elements must be evaluated in developing such procedures and to explore some possible avenues. Specifically, we have developed two monitoring tools using forecasted probabilities obtained from a multivariate logit model of banking crises. The first is an early-warning system that issues a signal when the probability of a forecasted crisis exceeds a certain threshold. The appropriate threshold for issuing a warning can be chosen based on the costs of missing a crisis and the benefits of avoiding a false alarm. The second monitoring tool is a system for rating bank fragility. Both monitoring tools can be used to economize on precautionary costs by pointing to cases of high fragility that warrant more in-depth monitoring.

Evaluating banking sector fragility along these lines is subject to several potential errors common to all exercises based on forecasts. First, the regression coefficients used to compute the forecasted probability of a crisis are only estimates of the true parameters. Second, new crises may be of a different nature than those experienced in the past, so that the coefficients derived from in-sample estimation may be of limited use out-of-sample. This problem may be particularly severe since banking crises tend to be rare events, and, even though the panel used for in-sample estimation is large (766 observations), crisis episodes only number 36.

Third, forecasts of the explanatory variables are likely to incorporate errors, as vividly illustrated by the example of the five recent Asian crises. Large forecast errors, in turn, may severely distort the assessment of fragility.[11] One way to reduce the impact of forecast errors is to develop alternative scenarios for the explanatory variables and to examine banking sector fragility in the context of such scenarios. This would be particularly useful, because in many cases banking crises are triggered by extreme behavior in one or more explanatory variables (a currency collapse, a bout of inflation, a drastic deterioration in the terms of trade) in a context in which other elements also contribute to overall fragility. Routine forecasts of economic variables rarely capture extreme events of this sort, which instead tend to be discussed as "risk elements" of the overall picture.[12] The frame-

11. One direction in which this work can be extended is to explore alternative model specifications and compare them from the point of view of their usefulness for forecasting (see, for instance, Diebold 1997). Here we have used a specification developed in our previous work after eliminating explanatory variables for which forecasts were not readily available. It could be that an even more parsimonious specification is more suitable for forecasting purposes. We leave this issue to future extensions.

12. This is certainly true of IMF forecasts, which often tend to be excessively optimistic (Mussa and Savastano 1999). For the Asian countries we computed crisis probabilities using the most pessimistic forecasts from the *Consensus Forecasts* group, but this did not lead to a substantial increase in forecasted crisis probabilities.

work developed here would lend itself easily to the evaluation of fragility in alternative scenarios, since it allows us to isolate the contribution of each explanatory variable to the forecasted crisis probability.

Another important caveat is that, although aggregate variables can convey information about the general economic conditions that tend to be associated with banking sector fragility, they are silent about the situation of individual banks or specific segments of the banking sector. So they would not detect crises that may develop from specific weaknesses in some market segments and spread through contagion. Also informed observers who are familiar with a particular country are likely to be in a better position to detect signs of incoming trouble, so the information generated by a quantitative approach such as ours should complement, not replace, other sources of information.

A final message from this exercise is that, to be useful, a monitoring system must be designed to fit the preferences of the decisionmaker. Thus the development of a system must be the outcome of an interactive process that involves both econometricians and policymakers.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Calvo, Guillermo A. 1996. "Capital Flows and Macroeconomic Management: Tequila Lessons." *International Journal of Finance and Economics* 1(3):207–24.

Caprio, Gerard, and Daniela Klingebiel. 1996. "Dealing with Bank Insolvencies: Cross-Country Experience." World Bank Policy Research Working Paper 1620. Washington, D.C.: World Bank.

Consensus Economics. Various years. *Consensus Forecasts*. London.

Demirgüç-Kunt, Aslı, and Enrica Detragiache. 1998. "The Determinants of Banking Crises in Developing and Developed Countries." IMF *Staff Papers* 45(1):81–109.

———. 1999. "Financial Liberalization and Financial Fragility." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998*. Washington, D.C.: World Bank.

Diebold, Francis. 1997. *Elements of Forecasting*. Cincinnati: South-Western College Publishing.

Eichengreen, Barry, and Andrew K. Rose. 1998. "Staying Afloat While the Wind Shifts: External Factors and Emerging-Market Banking Crises." NBER Working Paper 6370. National Bureau of Economic Research, Cambridge, Mass. Processed.

Financial Times. Various years. *Currency Forecaster*. Alexandria, Virginia.

Gavin, Michael, and Ricardo Hausman. 1996. "The Roots of Banking Crises: The Macroeconomic Context." In Ricardo Hausman and Liliana Rojas-Suárez, eds., *Volatile Capital Flows: Taming Their Impact on Latin America*. Washington, D.C.: Inter-American Development Bank.

Goldstein, Morris, and John Hawkins. 1998. "The Origins of the Asian Financial Turmoil." Reserve Bank of Australia, Sydney. Processed.

Hardy, Daniel, and Ceyla Pazarbasioglu. 1998. "Leading Indicators of Banking Crises: Was Asia Different?" IMF Working Paper 91. International Monetary Fund, Washington, D.C. Processed.

Honohan, Patrick. 1997. "Banking System Failures in Developing and Transition Countries: Diagnosis and Prediction." BIS Working Paper 39. Bank for International Settlements, Basel. Processed.

IMF (International Monetary Fund). 1997. "World Economic Outlook. Interim Assessment." Washington, D.C. Processed.

———. 1998. *World Economic Outlook*. Washington, D.C.

Kaminsky, Graciela L. 1998. "Currency and Banking Crises: The Early Warnings of Distress." International Finance Discussion Paper 629. Board of Governors of the Federal Reserve System, Washington, D.C.

Kaminsky, Graciela, and Carmen M. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance of Payments Problems." *American Economic Review* 89(3):473–500.

Lindgren, Carl J., Gillian García, and Michael Saal. 1996. *Bank Soundness and Macroeconomic Policy*. Washington, D.C.: International Monetary Fund.

Mishkin, Frederic S. 1996. "Understanding Financial Crises: A Developing Country Perspective." NBER Working Paper 5600. National Bureau of Economic Research, Cambridge, Mass.

Mussa, Michael, and Miguel Savastano. 1999. "The IMF Approach to Stabilization." *NBER Macroeconomic Annual* 14. Cambridge, Mass.: MIT Press.

Radelet, Steven, and Jeffrey Sachs. 1998. "The Onset of the Asian Financial Crises." In Paul Krugman, ed., *Currency Crises*. Chicago: University of Chicago Press.

Rojas-Suárez, Liliana. 1998. "Early Warning Indicators of Banking Crises: What Works for Developing Countries?" Research Department, Inter-American Development Bank, Washington, D.C. Processed.

Sachs, Jeffrey, Aaron Tornell, and Andres Velasco. 1996. "Financial Crises in Emerging Markets: The Lessons from 1995." *Brookings Papers on Economic Activity* 1(1996): 147–98.